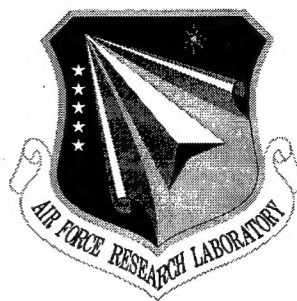


AFRL-SN-RS-TR-1999-185

Final Technical Report

August 1999



**PARAMETRIC AND MODEL BASED ADAP-
TIVE DETECTION ALGORITHMS FOR NON-
GAUSSIAN INTERFERENCE BACKGROUNDS**

ARCON Corporation

Muralidhar Rangaswamy

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

**AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK**

DTIC QUALITY INSPECTED 4

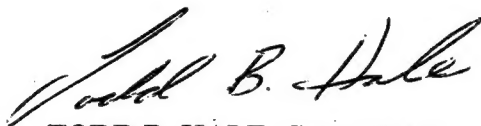
19991022 008

Although this report references a limited document (*), listed on page 43, no limited information has been extracted.

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-SN-RS-TR-1999-185 has been reviewed and is approved for publication.

APPROVED:



TODD B. HALE, Capt, USAF
Project Engineer

FOR THE DIRECTOR:



ROBERT G. POLCE, Acting Chief
Rome Operations Office
Sensors Directorate

If your address has changed or if you wish to be removed from the Air Force Research Laboratory Rome Research Site mailing list, or if the addressee is no longer employed by your organization, please notify AFRL/SNRT, 26 Electronic Pky, Rome, NY 13441-4514. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE August 1999	3. REPORT TYPE AND DATES COVERED Final Feb 97 - Oct 98		
4. TITLE AND SUBTITLE PARAMETRIC AND MODEL BASED ADAPTIVE DETECTION ALGORITHMS FOR NON-GAUSSIAN INTERFERENCE BACKGROUNDS		5. FUNDING NUMBERS C - F30602-97-C-0050 PE - 61102F PR - 2304 TA - E8 WU - PJ		
6. AUTHOR(S) Muralidhar Rangaswamy, Ph.D.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) ARCON Corporation 260 Bear Hill Rd Waltham MA 02451-1080		8. PERFORMING ORGANIZATION REPORT NUMBER N/A		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory/SNRT 26 Electronic Pky Rome Ny 13441-4514		10. SPONSORING/MONITORING AGENCY REPORT NUMBER AFRL-SN-RS-TR-1999-185		
11. SUPPLEMENTARY NOTES Air Force Research Laboratory Project Engineer: Capt Todd Hale, SNRT, (315)330-1896				
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release: Distribution unlimited.		12b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200 words) This report presents model based adaptive signal processing methods for target detection in a background of non-Gaussian interference. Several candidate algorithms are derived and important insights pertaining to their structure are documented in this report. Performance analysis of these algorithms is discussed in some detail. It is seen that the parametric adaptive matched filter (PAMF) offers the potential for significantly improved performance in non-Gaussian interference scenarios, while leading to considerably lower secondary data support requirements compared to classical adaptive processing methods. This is due to the use of a parametric method that employs a low model order to approximate the interference spectral characteristics. Another reduced rank adaptive algorithm considered in this study is the principal component inverse (PCI) method.				
14. SUBJECT TERMS Adaptive detection, parametric adaptive matched filter, principal components inverse, canonical correlations			15. NUMBER OF PAGES 64	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAR	

Contents

1	Introduction	1
1.1	Motivation	1
1.2	The Radar Environment	1
1.3	Report Organization and Overview of Results	4
1.4	Notation	4
1.5	Publications	4
2	Parametric Adaptive Matched Filter	6
2.1	Overview	6
2.2	Introduction	6
2.3	Problem Statement	7
2.4	The Parametric Detection Method	8
2.4.1	Maximum Likelihood Estimate of Signal Amplitude	9
2.4.2	Parametric Model Based Likelihood Ratio Test	9
2.5	Performance Analysis	15
2.6	Conclusions	19
3	A Unified Framework for STAP	21
3.1	Overview	21
3.2	Motivation	21
3.3	Problem Statement	22
3.4	Canonical Coordinates and Applications	24
3.5	Entropy and Canonical Coordinates	26
3.6	Conclusions	27
4	Adaptive Signal Processing for Non-Gaussian Clutter Back- grounds	28
4.1	Overview	28
4.2	Introduction	28
4.3	Problem Statement	29
4.4	Adaptive Processing Method	30
4.4.1	Maximum Likelihood Estimate of Covariance Matrix	30
4.4.2	Maximum Likelihood Estimate of Signal Amplitude	31
4.4.3	Adaptive Processor	32

4.5	Performance Analysis	33
4.6	Conclusions	36
5	Adaptive Signal Processing in Structured Interference Back-	
	grounds	37
5.1	Overview	37
5.2	Introduction	37
5.3	Problem Statement	39
5.4	The PCI Method	39
5.5	PCI for Non-Gaussian Interference	41
5.6	Conclusion	41
6	Conclusions	42

List of Figures

2.1	Detection Architecture of PAMF	13
2.2	Probability of Detection versus SINR: 1=Matched Filter in Gaussian Interference+White Gaussian Noise, 2=AMF in Gaussian Interference+White Gaussian Noise, 3,4=AMF in K-distributed Interference+ White Gaussian Noise	16
2.3	Probability of Detection versus SINR:1=MF in Gaussian Interference(Clutter+White Noise), 2=PAMF in Gaussian Interference, 3=AMF in Gaussian Interference	17
2.4	Probability of Detection versus SINR:1=PAMF in K-distributed clutter($\alpha = 0.1$)+WGN, 2=PAMF in K-distributed clutter($\alpha = 0.5$)+WGN, 3=AMF in Gaussian clutter+WGN, 4=AMF in K-distributed clutter($\alpha = 0.5$)+WGN, 5=AMF in K-distributed clutter($\alpha = 0.1$)+WGN	18
2.5	Probability of Detection versus SINR: 1=PAMF for K-distributed clutter ($\alpha = 0.1$)+WGN), 2=PAMF for K-distributed clutter ($\alpha = 0.5$) in K-distributed clutter ($\alpha = 0.1$)+WGN (Mismatched case)	19
4.1	Minimum Mean Squared Error Estimate of V^{-2}	34
4.2	Probability of Detection versus SINR K=10, J=1, N=5 $P_{fa} = 0.01$: 1=Adaptive Processor of eq(4.10) in K-distributed Noise $\alpha = 1$, $b = 1$, 2=Matched Filter in Gaussian Noise, 3=AMF in Gaussian Noise	35

Chapter 1

Introduction

1.1 Motivation

This work is motivated by the problem of adaptive radar target detection in clutter with unknown spectral and statistical characteristics. An important problem for the U.S. Air Force is adaptive radar signal processing with applications to airborne target detection, wide area surveillance and UAV platforms. These applications give rise to a number of open problems which are addressed under the framework of adaptive target detection in non-Gaussian interference in this effort. A key requirement in these applications is the maintenance of a constant false alarm probability.

1.2 The Radar Environment

In a typical radar system, the transmitter generates a coded waveform which is transmitted in free space. This waveform impinges on one or more targets and objects such as buildings, trees, water, and land depending on the operating environment. The reflections from a target constitute the desired signal, whereas reflections from other objects contribute to the clutter returns. The objective is to detect a desired signal embedded in unwanted signals. In practice, target detection is inhibited by a combination of clutter, one or more jammers, and background white noise.

The radar receiver front-end consists of an array of antenna elements receiving reflection from a source. The received reflection is an electromagnetic plane wave propagating in free space. The plane wave impinging on the array induces a voltage at each element of the array. The voltage induced at each array element forms the received data. Several snapshots of received data are collected either in the form of a data matrix or a concatenated data vector. In many instances radars operate at high pulse repetition frequencies, successive snapshots are highly correlated. Consequently, we need to account for the spatial and temporal correlation properties of the received data via a suitable model.

Typically, the spectral characteristics as well as the statistics underlying the clutter, jammer, and background noise are unknown. The following three important problems arise in this context.

1. Using the received data, it is necessary to determine whether or not a desired target is present. This issue motivates the problem of adaptive target detection.
2. The complex amplitude of the target embeds information pertaining to the radar target cross section, in addition to accounting for the dispersive effects of the propagating channel. Hence, the target complex amplitude needs to be estimated from the received data. This motivates the adaptive beamforming problem.
3. If operating in a hostile environment where one or more jammers are present, it is important to estimate the direction of arrival of the interfering sources with respect to the array normal. The direction of arrival estimation helps place nulls in the direction of the jammers. This issue motivates the adaptive direction of arrival estimation problem.

In this report, we concern ourselves with the problem of adaptive radar target detection using spatio-temporal data. In most instances, target detection is mainly inhibited by clutter. Consequently, the problem of adaptive target detection in clutter has received considerable attention [1,2]. In these reports the clutter statistics have been assumed to be Gaussian. A comprehensive treatment of STAP algorithms in Gaussian clutter for airborne radars is available in [1,2].

In many instances, clutter statistics deviate from the Gaussian assumption. For example, radar clutter observed from a high resolution radar looking down on the ground at low grazing angles is found to exhibit a probability density function with an extended tail [3]. The large-tailed PDF gives rise to high probability of false alarm. The Gaussian model for the clutter fails to predict the extended tail behavior. Consequently, non-Gaussian models have been proposed for the first order PDF of the clutter. Three commonly reported non-Gaussian PDF models are the Log-normal distribution, the Weibull distribution and the K-distribution. Furthermore the K-distribution has been found to satisfactorily account for the first order PDF of data from terrain and ocean scatter experiments [3,4]. The first two PDFs are based on empirical studies, whereas the K-distribution has physical significance in that the observed statistics can be related to the electromagnetic and geometric factors pertaining to the scattering surface [5].

More precisely, the probability density function (PDF) of ground clutter observed from airborne radars at low grazing angles exhibits a large tail [3]. The large tail gives rise to excessive false alarm probabilities. The Gaussian model for clutter fails to predict the large-tail behavior. Consequently non-Gaussian models are needed to describe the first order PDF of the clutter. Furthermore, in many instances radars employ a high pulse repetition frequency. Consequently,

successive clutter returns are highly correlated. Hence, in addition to the first order PDF, the pulse-to-pulse correlation of the clutter needs to be accounted for. When dealing with correlated non-Gaussian random variables, there is no unique model for their joint PDF. This is due to the arbitrarily large number of ubiquitous higher order dependencies. However, it becomes extremely difficult to measure the higher order dependencies in practice. On the other hand, it is possible to obtain information pertaining to the first order PDF and correlation function by using histogram fits to experimental data and power spectral density estimation or autocorrelation methods, respectively. Hence, a method for specifying the joint PDF of correlated non-Gaussian radar clutter based on the first order PDF and correlation function has considerable practical merit. Spherically invariant random processes (SIRP) provide a powerful mechanism for specifying the joint PDF of N correlated non-Gaussian random variables based on the first order PDF and correlation function [6, 7]. Therefore, the SIRP model for non-Gaussian clutter is used in all the studies undertaken in this work.

This work discusses methods for adaptive target detection in additive non-Gaussian interference which can be modeled by SIRPs. As a prelude to the target detection problem it is useful to discuss the target model in some detail. Typically, radar returns from a desired target undergo an attenuation and delay on account of the dispersive effects of the propagating medium. These effects are accounted via a complex amplitude. Additionally, the complex amplitude embeds information about the radar target cross-section. Therefore, the target is assumed to be known within a multiplicative complex amplitude of a known steering vector. The unknown complex amplitude can be modeled either as a deterministic parameter or as a random variable with each choice leading to a particular form for the adaptive detector. In this work, it is assumed the complex target amplitude is an unknown deterministic parameter.

The problem of adaptive target detection in non-Gaussian interference is then addressed. Important requirements in this problem include the reduction of training data support used for covariance estimation and the reduction of computational complexity of the algorithm. Classical methods for space-time adaptive processing (STAP) deal with target detection in Gaussian interference. These methods use the sample covariance matrix, which is the maximum likelihood estimate of the covariance matrix. However, the training data support requirement using the sample matrix approach is large. When a sample matrix based method is used for non-Gaussian interference cancellation, the training data requirements increase tremendously. Consequently, techniques for improving performance in non-Gaussian scenarios call for additional information about the interference. Accordingly, this report presents techniques which make use of additional information to reduce the training data and computational requirements.

1.3 Report Organization and Overview of Results

This report is organized as follows. Chapter 2 discusses the problem of adaptive target detection in non-Gaussian SIRPs using a multichannel model-based method. In this approach, the interference is approximated by a multichannel autoregressive (AR) process. More precisely, the interference is assumed to arise as the output of a multichannel system, whose transfer function is selected to give rise to the desired interference spectrum. The problem is then one of estimating the model coefficients (multichannel system transfer function) using training data. Significant savings in training data can be realized when the number of coefficients in the model (model order) is small. Low model order approximation has been found to satisfactorily account for the spectra of several simulated and real data scenarios. The resulting adaptive receiver is shown to be equivalent to a parametric adaptive matched filter compared to a data dependent threshold. Chapter 3 discusses the role of canonical correlations in unifying several recently proposed STAP algorithms. Chapter 4 addresses the problem of adaptive target detection in SIRPs using a covariance matrix based method. The resultant receiver structure is shown to be a generalized estimator-correlator. This provides several insights on the adaptive detection problem. In chapter 5, we discuss the principal component inverse (PCI) method and its application to SIRPs. Conclusions are presented in chapter 6.

1.4 Notation

1. Lowercase boldface symbols denote vectors.
2. Uppercase boldface symbols denote matrices.
3. Scalars are denoted by both upper-case and lower-case symbols.
4. $|\cdot|$ denotes the determinant of a matrix as well as the absolute value of a complex number. The argument of $|\cdot|$ specifies its usage.
5. The superscript $\hat{\cdot}$ denotes estimated quantities.

1.5 Publications

The following journal and conference publications resulted from this effort.

1. "A Parametric Detection Algorithm for Space-Time Adaptive Processing in Correlated Non-Gaussian Radar Clutter," M. Rangaswamy, J.H. Michels, to appear in the EURASIP Signal Processing journal.
2. "A parametric multichannel detection algorithm for correlated non-Gaussian random processes," M. Rangaswamy and J. H. Michels, IEEE national radar conference, Syracuse, NY, May 1997.

3. *"Space-Time Adaptive Processing (STAP) in Airborne Radar Applications,"* J. H. Michels, T. Tsao, B. Himed and M. Rangaswamy, IASTED international conference on signal processing and communications, Canary Islands, Spain, February 1998.
4. *"Performance of Principal Components Inverse (PCI) for Strong Low Rank Non-Gaussian Interference,"* B.E. Freburger, D.W. Tufts and M. Rangaswamy, invited paper at the conference on information sciences and systems, Princeton, NJ, March 1998.
5. *"Adaptive Signal Processing in Non-Gaussian Noise Backgrounds,"* M. Rangaswamy and J.H. Michels, proceedings of the ninth IEEE-SSAP Workshop, Portland, OR, September 1998.
6. *"A Unified Framework for Space-Time Adaptive Processing ,"* M. Rangaswamy, proceedings of the ninth IEEE-SSAP Workshop, Portland, OR, September 1998.

Chapter 2

Parametric Adaptive Matched Filter

2.1 Overview

This chapter discusses the problem of space-time adaptive processing for radar signal detection in additive correlated non-Gaussian clutter using a parametric model-based approach. The adaptive signal detection problem has been addressed extensively for the case of additive Gaussian clutter. However, the corresponding problem for the non-Gaussian case has received limited attention. The additive non-Gaussian clutter is assumed to be modeled by a spherically invariant random process (SIRP). The innovations based detection algorithm for the case of constant signal with unknown complex amplitude is derived. The resulting receiver structure is shown to be equivalent to a parametric adaptive matched filter compared to a data dependent threshold. Performance analysis of the derived receiver for the case of a K-distributed SIRV is presented.

2.2 Introduction

This investigation is motivated by a desire to detect signals in additive correlated non-Gaussian clutter using multichannel data. The problem of adaptive signal detection in additive noise background is of interest in several areas such as radar, sonar, and digital communications. This problem has been addressed in great detail when the background noise is Gaussian [8–13]. However, the corresponding problem for the case of additive, correlated non-Gaussian noise has received limited attention [14]. This is due to the lack of suitable models for describing the multivariate PDF of correlated non-Gaussian random variables. Spherically invariant random processes have been shown to provide an attractive mechanism for specifying the joint PDF of correlated non-Gaussian random variables [6, 7]. Sangston and Gerlach in [15] derived the SIRP model for non-

Gaussian radar clutter on the basis of a limit theorem arising from the physics of the scattering mechanism.

We concern ourselves with the adaptive detection problem in additive SIRP for a linear phased array radar, in which the received data can consist of returns from a target, clutter, jammers and white noise where the complex target amplitude and interference covariance matrix are unknown. In this chapter, interference consisting of clutter, jammer and white noise are denoted by "noise". The type of SIRP noise and the associated shape parameter values are assumed to be known a priori. This investigation generalizes the work of [14] in that the dispersive effects of the propagating medium are accounted for through an unknown complex amplitude. The adaptive processor is shown to be equivalent to a parametric adaptive matched filter compared to a data dependent threshold. Properties of the test and its performance analysis are presented.

This chapter is organized as follows. In section 2 we present the problem statement. Section 3 discusses the derivation of the innovations based receiver. Performance analysis of the derived method is discussed in section 4. Conclusions and directions of future research are presented in section 5.

2.3 Problem Statement

We outline the problem of interest in this section. N snapshots are observed using a J element narrowband uniform linear array with equally spaced elements. Using this data, we need to adaptively detect the presence of a desired signal. This is equivalent to the following statistical hypothesis testing problem in the presence of nuisance parameters.

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{y} \\ H_1 : \mathbf{x} &= a\mathbf{s} + \mathbf{y} \end{aligned} \quad (2.1)$$

where

$a \equiv$ Unknown complex signal amplitude.

$\mathbf{x} \equiv$ Received complex data vector under the two hypotheses.

$\mathbf{s} \equiv$ Desired $JN \times 1$ complex signal vector.

$\mathbf{y} \equiv JN \times 1$ complex SIRV with known characteristic PDF.

$H_0, H_1 \equiv$ The null and alternate hypothesis, respectively.

It is assumed that the $JN \times JN$ Hermitian covariance matrix of \mathbf{y} , denoted by Σ , is unknown.

We briefly state the following important properties of SIRVs and refer the interested reader to [6, 7, 14] and the references contained therein for further detail.

1. Every SIRV \mathbf{y} is equivalent to the product of a Gaussian random vector \mathbf{z} and an independent non-negative random variable V , with PDF $f_V(v)$, defined to be the characteristic PDF of the SIRV. Consequently, $\mathbf{y} = \mathbf{z}v$. The PDF of \mathbf{y} is given by

$$f_Y(\mathbf{y}) = \pi^{-JN} |\Sigma|^{-1} h_{2JN}(q) \quad (2.2)$$

where $q = \mathbf{y}\Sigma^{-1}\mathbf{y}$ and

$$h_{2JN}(q) = \int_0^\infty v^{-2JN} \exp(-\frac{q}{v^2}) f_V(v) dv \quad (2.3)$$

2. SIRVs are closed with respect to linear transformations. In other words every linear transformation on a SIRV results in another SIRV with the same characteristic PDF.
3. Minimum mean square error estimation (MMSE) problems involving SIRVs result in linear estimators.

An important issue in the hypothesis testing problem of eq (2.1) is the choice of the model for the complex signal amplitude. The complex signal amplitude, a , can be modeled either as a deterministic but unknown parameter or as a random variable. The latter choice does not permit the development of a uniformly most powerful (UMP) test even in the case of Gaussian noise and perforce precludes the development of a UMP test in the non-Gaussian case. Therefore, a is treated as a deterministic but unknown complex signal amplitude in this work.

2.4 The Parametric Detection Method

The hypothesis testing problem of eq (2.1) is now considered. No optimality property of the derived test is claimed. We consider the likelihood ratio, given a and Σ , shown by

$$\Lambda(\mathbf{x}|a, \Sigma) = \frac{h_{2JN}(q_1)}{h_{2JN}(q_0)} \quad (2.4)$$

where

$$\begin{aligned} q_0 &= \mathbf{x}^H \Sigma^{-1} \mathbf{x} \\ q_1 &= (\mathbf{x} - a\mathbf{s})^H \Sigma^{-1} (\mathbf{x} - a\mathbf{s}) \\ h_{2JN}(w) &= \int_0^\infty v^{-2JN} \exp(-\frac{w}{v^2}) f_V(v) dv \\ &\stackrel{H}{=} \text{Hermitian Transpose} \end{aligned} \quad (2.5)$$

The maximization of eq (2.4) over the nuisance parameters, a and Σ , is a difficult problem in general. No UMP test is available for this problem even in the Gaussian case ($h_{2JN}(q) = \exp(-q)$). Therefore, the approach used in this work consists of replacing the unknown parameters in eq (2.4) by their estimates. More precisely, we use a maximum likelihood estimate of a while employing a parametric estimate for Σ . The parametric estimate of Σ is optimal in a minimum mean square error sense. It has been shown in [16] that the parametric estimate for Σ is an approximate maximum likelihood solution in the Gaussian case. Specifically, we assume the SIRP is approximated by an autoregressive (AR) model. It is assumed that the order of the AR model is known a priori. In practice, the model order is unknown and must be determined for each scenario. This is a subject for future investigation.

2.4.1 Maximum Likelihood Estimate of Signal Amplitude

Let $\hat{\Sigma}$ denote the parametric estimate of Σ . The parametric estimate of Σ is discussed in some detail in [14]. We consider the maximization of eq (2.4) over a . Since a appears only in the numerator of eq (2.4), the maximum likelihood estimate of a is obtained from

$$\frac{\partial \Lambda(\mathbf{x}|a, \Sigma)}{\partial a} = \frac{\partial h_{2JN}(q_1)}{\partial a} = 0. \quad (2.6)$$

Note that $\frac{\partial h_{2JN}(q_1)}{\partial a} = \frac{\partial h_{2JN}(q_1)}{\partial q_1} \frac{\partial q_1}{\partial a}$. However, from eq (2.5), it follows that $h_{2JN}(\cdot)$ is a monotonically decreasing function. Consequently,

$$\frac{\partial h_{2JN}(q_1)}{\partial q_1} = - \int_0^\infty v^{-2JN-2} \exp(-\frac{q_1}{v^2}) f_V(v) dv < 0. \quad (2.7)$$

Hence, the desired solution is obtained by minimizing q_1 over a . We have

$$q_1 = (\mathbf{x} - a\mathbf{s})^H \hat{\Sigma}^{-1} (\mathbf{x} - a\mathbf{s}) = \|\hat{\Sigma}^{-\frac{1}{2}} (\mathbf{x} - a\mathbf{s})\|^2 \geq 0 \quad (2.8)$$

where $\|\cdot\|^2$ denotes the squared norm. From eq (2.8), q_1 is minimized when the squared norm is zero. This results when $a = \hat{a}$, where

$$\hat{a} = \frac{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}}. \quad (2.9)$$

Finally, we may rewrite q_1 as

$$\hat{q}_1 = \mathbf{x}^H \hat{\Sigma}^{-1} \mathbf{x} - \frac{|\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}} = \hat{q}_0 - \frac{|\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}}, \quad (2.10)$$

where $\hat{q}_0 = \mathbf{x}^H \hat{\Sigma}^{-1} \mathbf{x}$.

2.4.2 Parametric Model Based Likelihood Ratio Test

We derive the parametric model based likelihood ratio test by two different methods, called the block form and sequential form, respectively. The latter form is implemented in this work, while the former is helpful in discussing the properties of the test. Significantly, the test does not involve the explicit formation or inversion of a sample covariance matrix. Consequently, the algorithm is not restricted by the training data support of $K \geq JN$. In this method, the observed data processes are assumed to arise at the output of a multichannel linear system driven by white noise. Specifically, the observed data process is assumed to follow a multichannel autoregressive (AR) model. Given the transfer function of the multichannel linear system, the transfer function of the inverse filter which whitens the observed data is readily specified. The whitened data is termed as an innovations process. The innovations process is statistically equivalent to the observed data process. An excellent discussion of this issue can be found in [14] and the references therein. The transfer function of the

multichannel linear system and also that of the inverse filter are uniquely determined by specifying the multichannel AR model parameters. In the block form, these parameters are specified via the matrix coefficients of linear prediction and the diagonal error residual covariance matrix whereas in the sequential method these are denoted by the p^{th} order coefficients of linear prediction and the driving white noise covariance matrix respectively.

Block Form

We derive the parametric detection algorithm and show that the resulting test is equivalent to a parametric adaptive matched filter compared to a data-dependent threshold. The likelihood ratio of eq (2.4) is expressed as

$$\Lambda(\mathbf{x}|\hat{\mathbf{a}}, \hat{\Sigma}) = \frac{h_{2JN}(\hat{q}_1)}{h_{2JN}(\hat{q}_0)}. \quad (2.11)$$

We consider an \mathbf{LDL}^H decomposition of $\hat{\Sigma}$ with the \mathbf{L} and the \mathbf{D} matrices as defined in [14]. Following the method and notation of [14] we have

$$\begin{aligned} \hat{\Sigma} &= \mathbf{LDL}^H \\ \hat{\Sigma}^{-1} &= \mathbf{L}^{-1H} \mathbf{D}^{-1} \mathbf{L}^{-1} \\ \hat{q}_0 &= \mathbf{x}^H \mathbf{L}^{-1H} \mathbf{D}^{-1} \mathbf{L}^{-1} \mathbf{x} = \gamma^H \mathbf{D}^{-1} \gamma \\ \hat{q}_1 &= \gamma^H \mathbf{D}^{-1} \gamma - \frac{|\mathbf{s}_0^H \mathbf{D}^{-1} \gamma|^2}{\mathbf{s}_0^H \mathbf{D}^{-1} \mathbf{s}_0} \\ \gamma &= \mathbf{L}^{-1} \mathbf{x} \\ \mathbf{s}_0 &= \mathbf{L}^{-1} \mathbf{s} \end{aligned} \quad (2.12)$$

where γ is the innovations process defined in [14]. The innovations based likelihood ratio test is then expressed as

$$\Lambda\{\gamma\} = \frac{h_{2JN}(\hat{q}_1)}{h_{2JN}(\hat{q}_0)} \underset{H_0}{\overset{H_1}{>}} T \quad (2.13)$$

where T is a suitably chosen threshold depending on the desired false alarm probability. Eq (2.13) can be rewritten as

$$h_{2JN}(\hat{q}_1) \underset{H_0}{\overset{H_1}{>}} T h_{2JN}(\hat{q}_0). \quad (2.14)$$

Taking $h_{2JN}^{-1}(\cdot)$ of both sides of eq (2.14) gives

$$\hat{q}_1 \underset{H_0}{\overset{H_1}{>}} -h_{2JN}^{-1}[T h_{2JN}(\hat{q}_0)]. \quad (2.15)$$

The negative sign on the right hand side of eq (2.15) is due to the monotonically decreasing property of $h_{2JN}(\cdot)$. Using \hat{q}_1 from eq (2.12) in eq (2.15) and rearranging terms, the innovations based detection test takes the form

$$\frac{|\mathbf{s}_0^H \mathbf{D}^{-1} \gamma|^2}{\mathbf{s}_0^H \mathbf{D}^{-1} \mathbf{s}_0} \underset{H_0}{\overset{H_1}{>}} T_A \quad (2.16)$$

where $T_A = \hat{q}_0 + h_{2JN}^{-1}[Th_{2JN}(\hat{q}_0)]$. For the specific case of Gaussian noise, where $h_{2JN}(w) = \exp(-w)$, the test statistic of eq (2.13) is expressed as $\Lambda\{\gamma\} = \frac{h_{2JN}(\hat{q}_1)}{h_{2JN}(\hat{q}_0)} = \hat{q}_0 - \hat{q}_1 = \frac{|\mathbf{s}_0^H \mathbf{D}^{-1} \gamma|^2}{\mathbf{s}_0^H \mathbf{D}^{-1} \mathbf{s}_0}$ and the resulting test reduces to $\frac{|\mathbf{s}_0^H \mathbf{D}^{-1} \gamma|^2}{\mathbf{s}_0^H \mathbf{D}^{-1} \mathbf{s}_0} \underset{H_0}{\overset{H_1}{>}} T$ which is simply the parametric adaptive matched filter.

Recent work [12,13,17] examines the problem of adaptive signal detection in the framework of invariance with respect to rotation and gain transformations. The adaptive coherence estimator (ACE) test proposed therein takes the form

$$\frac{|\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s} \mathbf{x}^H \hat{\Sigma}^{-1} \mathbf{x}} \underset{H_0}{\overset{H_1}{>}} \eta \quad (2.17)$$

where $\hat{\Sigma} = \frac{1}{K} \sum_{i=1}^K \mathbf{x}_i \mathbf{x}_i^H$ and \mathbf{x}_i , $i = 1, 2, \dots, K$ denote the independent identically distributed training data. The ACE test of eq (2.17) has a pleasing geometry. It is also maximally invariant to rotation and gain transformations. More precisely, the same threshold setting is maintained when the test data \mathbf{x} is scaled by a gain factor G and the training data is scaled by a gain factor g . The CFAR feature of the test applies for the case of Gaussian interference. However, when the interference is a SIRV, each training data can be represented as $\mathbf{x}_i = \mathbf{z}_i \mathbf{V}_i$, $i = 1, 2, \dots, K$. If $\hat{\Sigma}$ is formed from this training data and used in the ACE test of eq (2.17), it is easy to show that the test statistic is no longer maximally invariant. This is due to the fact that different training data realizations are subject to different gain transformations. Consequently, the threshold is no longer invariant to the distribution of the scale factor. Hence, the CFAR feature is lost. Performance of this test in SIRVs is comparable to that of the AMF [10] operating in SIRVs and hence will not be considered further.

Sequential Form

Let $\mathbf{x}(m)$ denote one snapshot of the received data. The concatenated vector of snapshots is denoted by $\mathbf{x} = [\mathbf{x}(1)^T \mathbf{x}(2)^T \dots \mathbf{x}(N)^T]^T$. Each snapshot is a $J \times 1$ vector. Let $\mathbf{u}(n)$ denote the $J \times 1$ steering vector written in time series form. A multichannel one-step linear prediction error filter of order p with $J \times J$ matrix coefficients processes the data under the alternate hypotheses. Hence, \hat{q}_0 and \hat{q}_1 from eq (2.12) become

$$\begin{aligned} \hat{q}_0 &= \sum_{n=1}^N \gamma_0^H(n) \mathbf{D}_0^{-1} \gamma_0(n) \\ \hat{q}_1 &= \sum_{n=1}^N \gamma_0^H(n) \mathbf{D}_0^{-1} \gamma_0(n) - \frac{|\sum_{n=1}^N \mathbf{u}_0^H(n) \mathbf{D}_0^{-1} \gamma_0(n)|^2}{\sum_{n=1}^N \mathbf{u}_0^H(n) \mathbf{D}_0^{-1} \mathbf{u}_0(n)} \\ &= \sum_{n=1}^N \gamma_0^H(n) \mathbf{D}_0^{-1} \gamma_0(n) - |\hat{a}|^2 \sum_{n=1}^N \mathbf{u}_0^H(n) \mathbf{D}_0^{-1} \mathbf{u}_0(n) \end{aligned} \quad (2.18)$$

where $\gamma_0(n)$ is the whitened $J \times 1$ vector of samples for a given snapshot, \mathbf{D}_0 is the $J \times J$ diagonal covariance matrix of $\gamma_0(n)$ and $\mathbf{u}_0(n)$ is the transformed steering vector obtained by passing $\mathbf{u}(n)$ through the prediction error filter. Using \hat{q}_0 and \hat{q}_1 from eq (2.18) and following the algebra from eqs (2.13)-(2.16),

it follows that the test of eq (2.16) reduces to

$$\frac{|\sum_{n=1}^N \mathbf{u}_0^H(n) \mathbf{D}_0^{-1} \gamma_0(n)|^2}{\sum_{n=1}^N \mathbf{u}_0^H(n) \mathbf{D}_0^{-1} \mathbf{u}_0(n)} \underset{H_0}{\overset{H_1}{>}} T'_A \quad (2.19)$$

where $T'_A = \hat{q}_0 + h_{2JN}^{-1}[Th_{2JN}(\hat{q}_0)]$, with \hat{q}_0 being given by eq (2.18) and $\hat{a} = \frac{\sum_{n=1}^N \mathbf{u}_0^H(n) \mathbf{D}_0^{-1} \gamma_0(n)}{\sum_{n=1}^N \mathbf{u}_0^H(n) \mathbf{D}_0^{-1} \mathbf{u}_0(n)}$. The important difference between this derivation and the previous derivation is that only the p^{th} row of the \mathbf{L} matrix (in the previous derivation) is used here because we restrict ourselves to an AR filter of order p . Thus, we use the p^{th} order coefficients of linear prediction. The previous derivation makes use of the entire \mathbf{L} matrix which contains the coefficients of linear prediction of orders 1 to p [18]. However, as with the test of eq (2.16), the test of eq (2.19) also involves a data dependent threshold, which cannot be determined without a priori knowledge of $f_V(v)$. Consequently, the detection test implemented here is given by

$$\ln[\Lambda\{\gamma\}] = \ln[h_{2JN}(\hat{q}_1)] - \ln[h_{2JN}(\hat{q}_0)] \underset{H_0}{\overset{H_1}{>}} \ln(T) \quad (2.20)$$

where T is a suitably chosen threshold depending on the desired false alarm probability and \hat{q}_0, \hat{q}_1 are specified in eq (2.18). A block diagram of the test implementation is shown in Figures 2.1(a) and 2.1(b). The two figures are equivalent except that Figure 2.1(b) explicitly denotes the estimation of a .

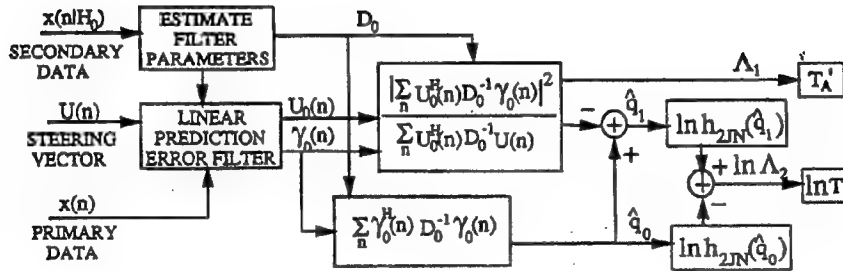
Properties Of The Parametric Model Based Test

From eq (2.16) it is clear that the innovations based detection test of eq (2.13) is equivalent to a parametric adaptive matched filter compared to a data dependent threshold. The test of eq (2.16) has several interesting properties.

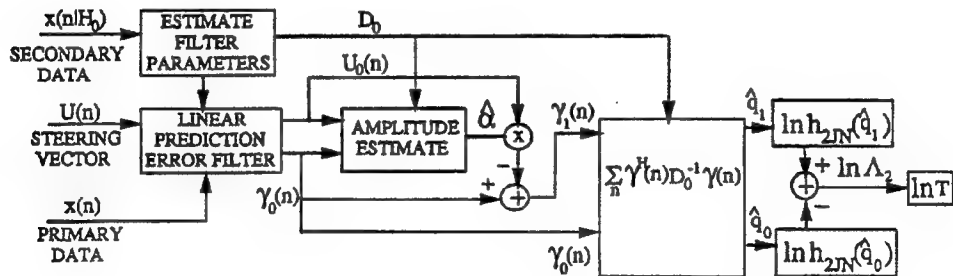
1. The test statistic of eq (2.16) is simply the squared magnitude of the inner product of a transformed signal vector and the innovations vector, introduced by the matrix \mathbf{D}^{-1} normalized by the norm of the transformed signal vector. Thus, the innovations based detection method can be cast in the framework of $|\mathbf{w}^H \mathbf{x}|^2 \underset{H_0}{\overset{H_1}{>}} T_A$ where $\mathbf{w} = \frac{\mathbf{D}^{-1} \mathbf{s}_0}{\sqrt{\mathbf{s}_0^H \mathbf{D}^{-1} \mathbf{s}_0}}$ is the normalized projection of the transformed desired signal vector onto the null space of the noise covariance matrix.
2. We consider the case when $h_{2JN}(w) = \exp(-w)$ (Gaussian noise). Then, the test of eq (2.16) reduces to

$$\frac{|\mathbf{s}_0^H \mathbf{D}^{-1} \gamma|^2}{\mathbf{s}_0^H \mathbf{D}^{-1} \mathbf{s}_0} \underset{H_0}{\overset{H_1}{>}} T' \quad (2.21)$$

where $T' = \ln(T)$. In this case, note that the data dependence of the threshold vanishes and the test becomes the parametric version of the



a.



b.

Figure 2.1: Detection Architecture of PAMF

classical adaptive matched filter (AMF) test of [10]. Thus, we conclude that the innovations based detection algorithm (IBDA) for SIRPs is a generalization of the AMF derived in [10].

3. The test of the form of eq (2.16) provides further insight into the detection results of [14]. Comparison of curves 1 and 2 of Fig 2 in [14] reveals that the IBDA for non-Gaussian noise backgrounds has the potential for significant performance improvement at low signal-to-noise-ratio (SNR). Also Fig 2 of [14] revealed that at large SNR (> 5 dB), the receiver for the case of Gaussian noise outperforms the IBDA for non-Gaussian backgrounds. This performance inversion can be explained by casting the IBDA of [14] in the form of eq (2.16). Following the method presented in the derivation of eq (2.16), the IBDA of [14] can be expressed as

$$\mathcal{Re}[\mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{x}] \underset{H_0}{\overset{H_1}{>}} \frac{1}{2} \mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{s} + \frac{1}{2} q_0 + \frac{1}{2} h_{2JN}^{-1}[Th_{2JN}(q_0)] \quad (2.22)$$

$$q_0 = \mathbf{x}^H \boldsymbol{\Sigma}^{-1} \mathbf{x}.$$

For the special case of Gaussian noise eq (2.22) reduces to

$$\mathcal{Re}[\mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{x}] \underset{H_0}{\overset{H_1}{>}} \frac{1}{2} \mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{s} + \frac{\eta}{2} \quad (2.23)$$

We first recognize that eqs (2.22) and (2.23) correspond to a matched filter statistic. In Fig 2 of [14], the noise was assumed to be white. Hence, $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$. Furthermore, since $\mathbf{s}^H \mathbf{s} = E$ (constant energy), the SNR is controlled by σ^2 . Specifically, low SNR corresponds to large values of σ^2 and vice versa. Large values of σ^2 (low SNR) result in a reduction in the contributions of the first two terms of the threshold in eq (2.22). We then examine the contribution of the third term in the threshold of eq (2.22). Since $h_{2JN}(\cdot)$ is a monotonically decreasing function, large values of σ^2 result in small values of q_0 and hence large values of $h_{2JN}(\cdot)$. This results in a large value for the argument of $h_{2JN}^{-1}(\cdot)$. Again due to the monotonicity of $h_{2JN}(\cdot)$, it follows that $h_{2JN}^{-1}(\cdot)$ gives rise to a large negative contribution, thereby resulting in a reduction of the threshold. This additional reduction in the threshold arising from the third term of eq (2.22) is responsible for improved detection performance at low SNR. The Gaussian case results in a constant threshold which is greater than that of the non-Gaussian IBDA. Therefore, its performance at low SNRs suffers.

However, at high SNRs (small values of σ^2), the first two terms of eq (2.22) give rise to large contributions while the third term yields a very small negative contribution. Hence, the threshold is larger than that for the Gaussian case. Therefore, the Gaussian receiver outperforms the non-Gaussian IBDA.

4. A possible criticism against the test of eq (2.16) and eq (2.19) is that the threshold is data dependent. Thus, determination of the data dependent threshold requires a priori knowledge of the specific type of SIRP noise background, which is generally unknown. Furthermore, threshold specification involves calculation of $h_{2JN}^{-1}(\cdot)$. In many instances, the functional form of $h_{2JN}(\cdot)$ is extremely complicated. Therefore, it may not be possible to calculate $h_{2JN}^{-1}(\cdot)$ analytically. In such instances, numerical techniques must be used. These issues are currently under investigation. Therefore, in this work, we use the test of eq (2.20). It is extremely difficult to determine the PDF of the test statistic of eq (2.13) in closed form. Consequently, it is not possible to obtain closed form or near closed form results for the detection and false alarm probabilities. Hence, performance analysis must be carried out by Monte-Carlo simulations following the method of [14].
5. In general, $T > 1$. Consequently, $T_A > 0$. Also, since $h_{2JN}(\cdot)$ is a monotonically decreasing function, it follows that $T_A < \hat{q}_0$. Hence, $0 < T_A < \hat{q}_0$. The data dependence of the threshold vanishes for the case of Gaussian noise ($h_{2JN}(w) = \exp(-w)$). Another case where the data dependence vanishes is when $T = 1$. This case arises frequently in digital communication problems where a Bayesian hypothesis testing method involving cost functions is used. Then, although $h_{2JN}(w) \neq \exp(-w)$, the test of eq (2.16) reduces to that for the case of Gaussian noise.

2.5 Performance Analysis

Performance analysis of the detection architecture of Fig 2.1 is discussed for the case of a K-distributed SIRV. The K-distribution has been proposed as a model for the amplitude statistics of ground and sea clutter observed from high resolution airborne radars operating at low grazing angles. The K-distributed envelope PDF is useful for modeling non-Gaussian radar clutter [14] and is given by

$$f_R(r) = \frac{2b}{\Gamma(\alpha)} \left(\frac{br}{2}\right)^\alpha K_{\alpha-1}(br) \quad 0 \leq r \leq \infty \quad (2.24)$$

where α is the shape parameter of the distribution, b denotes the scale parameter of the distribution, and $K_N(\cdot)$ is the N^{th} order modified Bessel function of the second kind. The characteristic PDF for the K-distributed SIRV and the corresponding $h_{2JN}(q)$ are

$$\begin{aligned} f_V(v) &= \frac{2b}{\Gamma(\alpha)} (bv)^{2\alpha-1} \exp(-b^2 v^2) \quad 0 \leq v \leq \infty. \\ h_{2JN}(q) &= \frac{2b^{2JN}}{\Gamma(\alpha)} (b\sqrt{q})^{\alpha-JN} K_{\alpha-JN}(2b\sqrt{q}). \end{aligned} \quad (2.25)$$

Performance evaluation is carried out using Monte-Carlo simulation. For a false alarm probability, P_{fa} , of 0.01, the Monte-Carlo method uses $N_R = 10,000$ independent realizations of observation data for each detection run. Each detection run used one estimate of the sample covariance matrix (or AR parameters

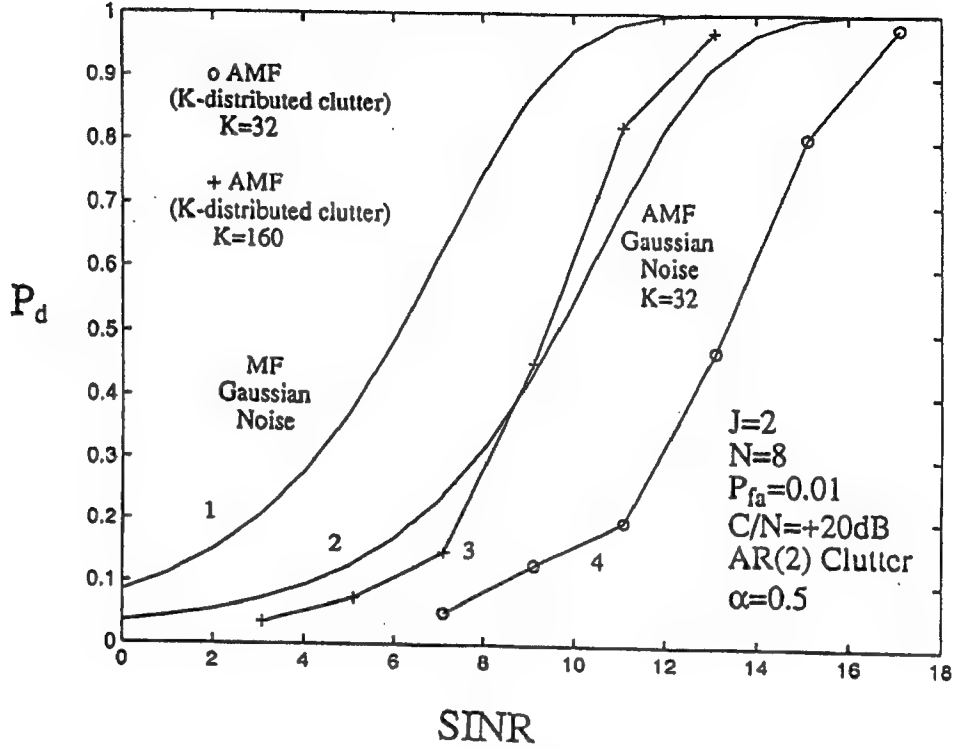


Figure 2.2: Probability of Detection versus SINR: 1=Matched Filter in Gaussian Interference+White Gaussian Noise, 2=AMF in Gaussian Interference+White Gaussian Noise, 3,4=AMF in K-distributed Interference+ White Gaussian Noise

for the parametric algorithm) based on a sample size of K realizations. The procedure is repeated for a total of $N_d = 1000$ detection runs.

The resulting detection probability, P_d , is compared to that of the classical adaptive matched filter of [10]. Figures 2.2-2.5 show the plot of P_d versus signal-to-interference plus noise ratio (SINR) for the case of a constant signal with unknown amplitude in partially correlated clutter [AR(2) model] plus white Gaussian noise. Detection results are based on two channels ($J=2$), eight pulses ($N=8$), and $P_{fa} = 0.01$ with sample support size (K) as a parameter. In Figure 2.2 curve 1 shows the performance of the matched filter (MF) receiver in Gaussian clutter plus white noise. Curve 2 depicts the performance of the adaptive matched filter (AMF) receiver in Gaussian clutter plus white noise with $K=32$. Curve 4 shows the degradation in performance of the AMF when the clutter is K-distributed with shape parameter $\alpha = 0.5$. The performance reduction is due to poor covariance estimator performance and the increased threshold setting due to the heavy tails of the K-distributed clutter process. In [19], analytical expressions are derived for the increase in sample support size needed to reduce the estimator error variance of the sample covariance

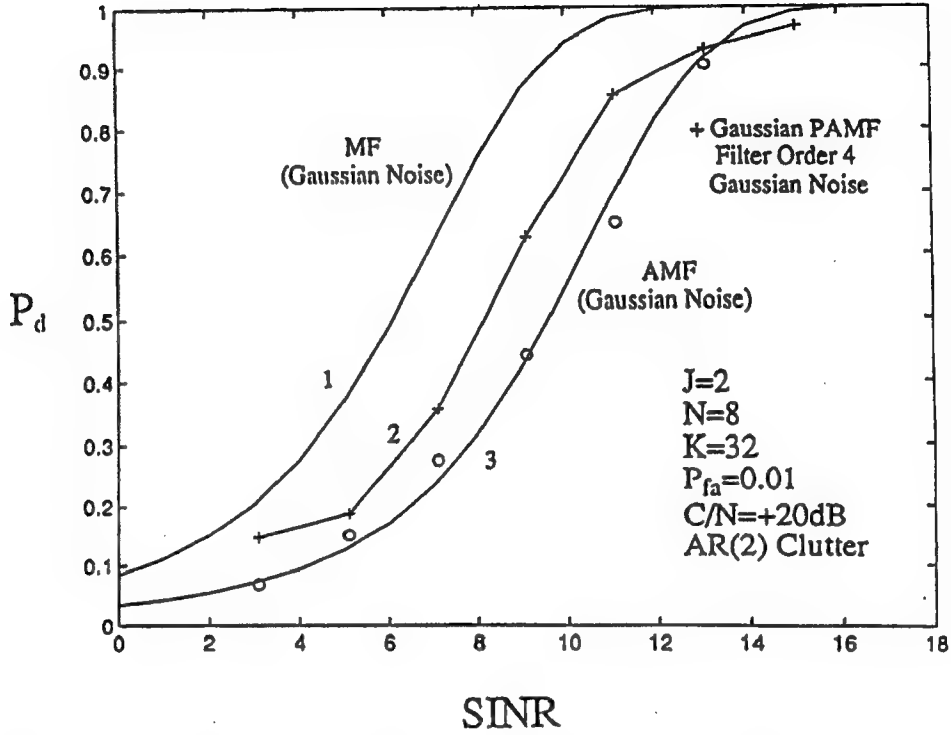


Figure 2.3: Probability of Detection versus SINR: 1=MF in Gaussian Interference(Clutter+White Noise), 2=PAMF in Gaussian Interference, 3=AMF in Gaussian Interference

estimator for SIRP clutter compared to that for the Gaussian case. For K -distributed processes, the increase is given by the factor $1 + \frac{2}{\alpha}$. Thus, for $\alpha = 0.5$, the sample support size needed for the sample covariance estimator for K -distributed SIRPs is five times greater than that for the Gaussian case to produce the same estimator error variance. In a recent effort [20], we present the performance improvement that can be obtained by using the maximum likelihood estimate of the covariance matrix for SIRPs. Use of this estimator will restore the maximal invariance and CFAR properties of the test of [12, 13] for SIRVs. Curve 3 shows the resulting performance improvement of the AMF in K -distributed clutter plus white Gaussian noise when the sample support size increased by a factor of five.

Figure 2.3 again shows P_d versus SINR for the MF and the AMF for Gaussian clutter plus white noise in curves 1 and 3, respectively. Monte-Carlo P_d results for the AMF are shown by the points designated 'o' for validation purposes. Curve 2 shows the Monte-Carlo P_d results (+) for the parametric adaptive matched filter (PAMF) designed for Gaussian clutter plus white noise. A matrix prediction error filter of order 4 was used. The Strand-Nuttall algorithm [21, 22]

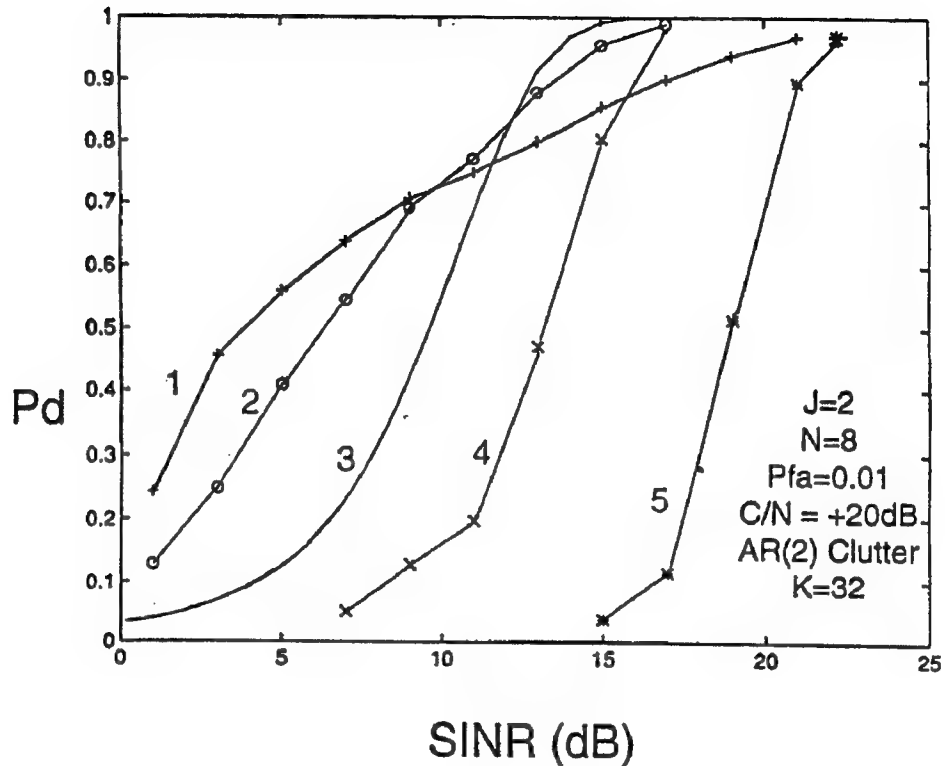


Figure 2.4: Probability of Detection versus SINR: 1=PAMF in K-distributed clutter($\alpha = 0.1$)+WGN, 2=PAMF in K-distributed clutter($\alpha = 0.5$)+WGN, 3=AMF in Gaussian clutter+WGN, 4=AMF in K-distributed clutter($\alpha = 0.5$)+WGN, 5=AMF in K-distributed clutter($\alpha = 0.1$)+WGN

was used to estimate the filter coefficients. Note the improved performance of the PAMF over the AMF.

In Figure 2.4, we first consider curves 3, 4, and 5, corresponding to the performance of the AMF in K-distributed clutter plus white Gaussian noise(WGN) with shape parameters $\alpha = \infty$ (Gaussian), $\alpha = 0.5$, and $\alpha = 0.1$, respectively. We observe a degradation in performance as the statistics of the clutter process deviate from Gaussianity. Curves 1 and 2 show the performance of the PAMF designed for K-distributed clutter process with shape parameters $\alpha = 0.5$ and $\alpha = 0.1$, operating in K-distributed clutter plus white noise for $\alpha = 0.5$ and $\alpha = 0.1$, respectively. These curves reveal several interesting features. First, we note that operation in non-Gaussian clutter offers the potential for significant performance improvement over the Gaussian case for a considerable range of SINR when the receiver is properly designed. Next, the convergence of curves 4 and 5 to curves 2 and 1, respectively reveals that as the SINR increases, the best achievable performance corresponds to the Gaussian receiver.

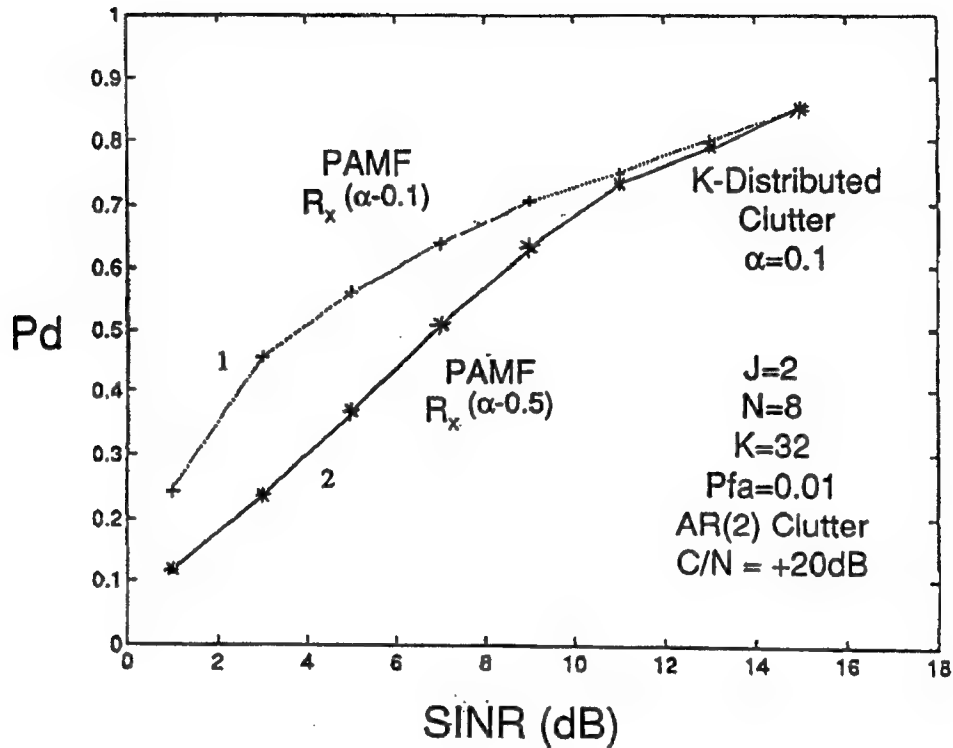


Figure 2.5: Probability of Detection versus SINR: 1=PAMF for K-distributed clutter ($\alpha = 0.1$)+WGN), 2=PAMF for K-distributed clutter ($\alpha = 0.5$) in K-distributed clutter ($\alpha = 0.1$)+WGN (Mismatched case)

Finally, in Figure 2.5 we show the degradation in performance at low SINR values when the non-Gaussian PAMF is mismatched to the clutter PDF. Curve 1 shows the performance of the PAMF designed for $\alpha = 0.1$ operating in K-distributed clutter plus white Gaussian noise with $\alpha = 0.1$ while curve 2 shows the performance of the same receiver operating in K-distributed clutter plus white Gaussian noise with $\alpha = 0.5$. The resulting performance degradation is higher than that reported in [14] for the case of a known signal amplitude.

2.6 Conclusions

This chapter addressed the problem of adaptive signal detection in the presence of additive SIRP noise using a parametric model-based approach. The derived receiver was shown to be equivalent to a parametric adaptive matched filter compared to a data dependent threshold. For the particular case of Gaussian noise, the receiver reduced to the parametric adaptive matched filter. Performance analysis of the PAMF for the case of K-distributed SIRP was presented. The

PAMF reveals the potential for significant performance improvement over the classical AMF when operating in non-Gaussian clutter backgrounds, especially at low SINR. It was also seen that the performance of the classical AMF degrades severely in non-Gaussian clutter backgrounds. Also the sample support size needed to improve the performance of the AMF in non-Gaussian clutter backgrounds becomes excessively large as the statistics of the clutter process deviate from Gaussianity.

Future work in this area must include an investigation of the properties of the data dependent threshold, CFAR features of the PAMF, robustness of the PAMF receiver as well as a procedure to determine the model order used for approximating the clutter spectral characteristics.

Chapter 3

A Unified Framework for STAP

3.1 Overview

This work provides a common framework for three recently proposed space-time adaptive processing (STAP) methods. A common goal of these methods is to reduce computational complexity and sample support requirements. It is shown that the canonical correlations model provides a mechanism for treating the STAP methods in a unified framework.

3.2 Motivation

This work is motivated by the adaptive signal processing problem in additive interference with unknown spectral characteristics. Applebaum [23] first studied this problem for adaptive antenna arrays using a feedback loop implementation of the adaptive processor. However, the technique was slow to converge to the steady-state solution. Space-time adaptive processing (STAP) for radar was first proposed by Brennan and Reed in [8, 24]. The sample matrix inversion method of [8] exhibited significantly improved convergence over the method of [23]. However, the work of [8] lacked the constant false alarm rate (CFAR) feature. The CFAR adaptive matched filter (AMF), a variant of the beamformer of [8] was independently derived by Chen [9] and Robey [10], respectively.

This work considers reduction of the computational complexity of STAP algorithms in interference scenarios with unknown covariance matrices. The analysis herein provides a unified framework for treating several candidate STAP algorithms for Gaussian and non-Gaussian interference statistics. Important considerations in this context are the sample support requirements for inferring the spectral characteristics of the interference and the computational complexity in forming the weight vector. The latter issue has been the focus of recent

studies. Goldstein [25] proposed the cross-spectral metric (CSM) for reducing the interference space dimension and for faster adaptive algorithm convergence. Tufts in [26, 27] developed the principal components inverse (PCI) method for the same problem. In [28], Rangaswamy and Michels derived the parametric adaptive matched filter (PAMF). While the analysis of [25]- [27] treats computational complexity for Gaussian interference, the analysis of [28] addresses both computational complexity as well as sample support requirements for Gaussian and non-Gaussian interference scenarios. Model-order determination, an important aspect of the PAMF, is a subject of ongoing investigation. Pados in [29] proposed the auxiliary vector STAP algorithm for reducing computational complexity and sample support. In a related effort, Roman [30] developed the adaptive sidelobe canceller using the canonical correlations model (CCM).

A unified framework for treating the PAMF, PCI, CSM, and auxiliary vector algorithms results from using the canonical correlations model (CCM). The analysis is carried out for known covariance matrix. Necessary modifications for adaptive estimation of the covariance matrix are then pointed out. Since this work relies on the methods of [25]- [27], we restrict ourselves to the case of Gaussian interference. However, in many instances non-Gaussian statistics have been reported for radar clutter returns from terrain and ocean [3, 4]. Spherically invariant random processes (SIRP) have been used extensively as models for correlated non-Gaussian clutter [6, 7]. Therefore, our future work will extend the analysis of this work to the case of SIRPs.

3.3 Problem Statement

The problem considered in [25]- [29] is one of target detection in the presence of nuisance parameters which is posed as the following statistical hypothesis testing problem:

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{y} \\ H_1 : \mathbf{x} &= a\mathbf{s} + \mathbf{y} \end{aligned} \quad (3.1)$$

where $\mathbf{x} \equiv$ received $JN \times 1$ complex observation vector, $\mathbf{s} \equiv$ known steering vector, $a \equiv$ unknown complex signal amplitude, and $\mathbf{y} \equiv$ complex Gaussian random vector with known Hermitian positive definite covariance matrix, \mathbf{R} . This is a classical signal detection problem in which the signal can be expressed as a linear combination of modes or basis vectors [31]. However, the modal coefficients are unknown. More precisely, the steering vector is the modal vector and the unknown complex signal amplitude is the modal coefficient. This association leads to the matched subspace filter interpretation of the matched filter discussed in [25]. It also provides the motivation for reducing the dimensionality and thus reduce the computational complexity of the algorithm.

The maximum likelihood (ML) estimate of a is given by

$$\hat{a} = \frac{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{x}}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}} \quad (3.2)$$

This estimate is the output of a minimum variance distortionless response (MVDR) beamformer given by $y = \mathbf{w}^H \mathbf{x}$ where $\mathbf{w} = \frac{\mathbf{R}^{-1} \mathbf{s}}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}}$. The beamformer output noise variance is given by $P = E[|\mathbf{w}^H \mathbf{x}|^2] = \frac{1}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}}$.

The ML estimate is then used in a likelihood ratio test, which takes on the form

$$\Lambda(\mathbf{x}) = \frac{|\mathbf{s}^H \mathbf{R}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}} \underset{H_0}{\overset{H_1}{>}} T \quad (3.3)$$

The test statistic of eq (3.3), which has several interesting interpretations, is called the direct form [25]. The first interpretation is that $\Lambda(\mathbf{x}) = \frac{|y|^2}{P}$ where y and P have been defined for the MVDR beamformer. This gives the interpretation that $\Lambda(\mathbf{x})$ is simply the MVDR beamformer output signal-to-noise-ratio (SNR). It is also evident that the MVDR beamformer weight vector is a normalized projection of the steering vector onto the null space of the noise covariance matrix. Since the interference is Gaussian, probability of detection is a monotonic function of the SNR. Therefore, maximizing SNR is equivalent to maximizing the probability of detection. Hence, the interpretation of the weight vector as a projection onto the null space of the noise covariance matrix is an elegant statement of SNR maximization.

A second interpretation is that of equivalence to a known level scalar CFAR processor. A scalar CFAR processor calculates the squared magnitude of the observed data normalized by the noise variance and compares to a pre-determined threshold. With reference to the test of eq (3.3), the scalar data is the output of the beamformer of [8] given by $\mathbf{s}^H \mathbf{R}^{-1} \mathbf{x}$ with output noise variance $\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}$. Since the threshold is independent of the actual noise covariance matrix, the test has an embedded CFAR feature which becomes clear from the analogy to a scalar CFAR processor.

A third interpretation is that $\Lambda(\mathbf{x})$ is a normalized squared magnitude correlator output (inner product). This may also be viewed as the result of maximizing a likelihood function with respect to an unknown complex amplitude and known covariance matrix. This interpretation of the test statistic provides insights into the CFAR property and the relationship to magnitude squared coherence between the vectors \mathbf{s} and \mathbf{x} . In an adaptive implementation, the work of [10] presents the maximization principle while using an estimated covariance matrix (sample covariance matrix).

A fourth interpretation is in terms of a vector subspace approach. This gives rise to the indirect form or sidelobe canceller form discussed in [25]. Note that the test statistic can be interpreted as a two stage operation of matched filtering followed by interference removal. The matched filtering consists of projecting the received observation vector onto the signal subspace. In this case, the signal subspace dimension is unity because of the single mode (steering vector). However, since the received data could consist of signal plus interference, the received data projection onto the signal subspace contains signal as well as interference components. Therefore, matched filtering is followed by interference removal. Interference removal is accomplished by estimating the interference component

in the signal subspace using the projection of the received data onto the orthogonal complement space or the noise subspace. The methods of [25]- [27] can be cast in this framework.

More precisely, these methods involve an eigen-decomposition of the interference covariance matrix. The PCI technique of [26, 27] retains the dominant eigenvalues (and the corresponding eigenvectors) of the interference covariance matrix. This leads to a reduced-dimension problem. Since the PCI method assumes that the interference covariance matrix is well approximated by a small number of strong low rank interferers plus white noise, best performance is attained when this assumption is satisfied. This is equivalent to the model order selection problem encountered in the PAMF of [28]. PCI performance degrades when a large number of dominant eigenvalues must be considered. The CSM however, considers a slightly different criterion for reducing dimensionality. In particular, the CSM calculates the squared magnitude of the steering vector projection along the eigenvectors of the noise covariance matrix normalized by the corresponding eigenvalues, and purges the sub-dominant normalized projections. This gives rise to a maximum SNR criterion or alternatively minimization of mean squared error in estimating the interference component in the signal subspace from projections in the orthogonal complement noise subspace. The signal blocking matrix, \mathbf{B} , in the sidelobe canceller form of implementation in [25] provides a set of orthonormal basis vectors for the orthogonal complement noise subspace. In the next section, we present some aspects of coherence and canonical coordinates and show that the method of [25] is an application of the canonical correlations model.

3.4 Canonical Coordinates and Applications

In problems of statistical hypothesis testing and parameter estimation, a compact representation of the information contained in the observed data is sought. Such a representation greatly aids in dimensionality and storage requirement reduction. In this study, we concern ourselves with the problem of minimum mean squared error estimation and the role of canonical correlations in this context. More precisely, in many problems of engineering importance, it is necessary to obtain information about a source vector given an observation vector. For example, the source vector could be a message signal generated at a transmitter and the observed data could represent the received signal. Signal propagation through a channel results in a received signal affected by noise and channel effects such as attenuation and group delay. In radar and sonar, the source vector is a steering vector. The transmitted signal is an electromagnetic (or acoustic wave) propagating in free space (or underwater). Due to the characteristics of the propagating medium, the transmitted signal undergoes attenuation and delay. In addition, the signal is contaminated by noise, clutter and jammer. Therefore, in these applications it becomes important to determine the similarity between the received signal and a replica of the transmitted signal (steering vector). Recently, we became aware of the related work of [32]. This

work contains the formulas for prediction error and entropy, which we derived independently.

The coherence between two vectors provides a measure of similarity between them [33]. Some mathematical details pertaining to coherence are in order. Let $\mathbf{z} \equiv l \times 1$ Source vector, $\mathbf{p} \equiv M - l \times 1$ Observed data vector, $\mathbf{R}_z \equiv$ Covariance matrix of \mathbf{z} , $\mathbf{R}_p \equiv$ Covariance matrix of \mathbf{p} , and $\mathbf{R}_{zp} \equiv$ Cross Covariance matrix between \mathbf{z} and \mathbf{p} . For convenience, the vectors \mathbf{z} and \mathbf{p} are assumed to have zero mean. The coherence matrix between \mathbf{z} and \mathbf{p} is given by

$$\mathbf{C}_{zp} = E[\mathbf{R}_z^{-\frac{1}{2}} \mathbf{z} (\mathbf{R}_p^{-\frac{1}{2}} \mathbf{p})^H] = \mathbf{R}_z^{-\frac{1}{2}} \mathbf{R}_{zp} \mathbf{R}_p^{-\frac{1}{2}} \quad (3.4)$$

The minimum mean squared error estimate of \mathbf{z} given \mathbf{p} is

$$\hat{\mathbf{z}} = E[\mathbf{z}|\mathbf{p}] = \mathbf{R}_{zp} \mathbf{R}_p^{-1} \mathbf{p} \quad (3.5)$$

The estimation error covariance matrix is given by

$$\mathcal{E} = E[(\mathbf{z} - \hat{\mathbf{z}})(\mathbf{z} - \hat{\mathbf{z}})^H] = \mathbf{R}_z - \mathbf{R}_{zp} \mathbf{R}_p^{-1} \mathbf{R}_{pz} \quad (3.6)$$

Using the definition of the coherence matrix, the error covariance matrix can be expressed as

$$\mathcal{E} = E[(\mathbf{z} - \hat{\mathbf{z}})(\mathbf{z} - \hat{\mathbf{z}})^H] = \mathbf{R}_z^{\frac{1}{2}} (\mathbf{I} - \mathbf{C}_{zp} \mathbf{C}_{zp}^H) \mathbf{R}_z^{\frac{1}{2}} \quad (3.7)$$

The canonical coordinate representation relies on a singular value decomposition (SVD) of the coherence matrix. More precisely, the coherence matrix can be expressed as

$$\mathbf{C}_{zp} = \mathbf{F} \mathbf{K} \mathbf{G}^H \quad (3.8)$$

where \mathbf{F} and \mathbf{G} denote the matrices whose columns are orthonormal eigenvectors of $\mathbf{C}_{zp} \mathbf{C}_{zp}^H$ and $\mathbf{C}_{zp}^H \mathbf{C}_{zp}$, respectively. \mathbf{K} is the diagonal matrix of singular values of \mathbf{C}_{zp} . The elements of \mathbf{K} represent the canonical correlations. The error covariance matrix can be expressed as

$$\mathcal{E} = \mathbf{R}_z^{\frac{1}{2}} (\mathbf{I} - \mathbf{F} \mathbf{K}^2 \mathbf{F}^H) \mathbf{R}_z^{\frac{1}{2}} \quad (3.9)$$

Each element of \mathbf{K} is the direction cosine between a unit variance random variable obtained from the canonical source coordinates and another unit variance random variable from the canonical measurement coordinates. The diagonal form of \mathbf{K} ensures that the i^{th} unit variance random variable from the canonical source coordinate is correlated only with the i^{th} unit variance random variable from the canonical measurement coordinate. Retaining the dominant direction cosines enables dimensionality reduction.

With reference to [25], and following the notation therein, $d = \mathbf{s}^H \mathbf{x}$, $\mathbf{b} = \mathbf{B} \mathbf{x}$, $\sigma_d^2 = E[|d|^2] = \mathbf{s}^H \mathbf{R}_s \mathbf{s}$, $\mathbf{R}_b = \mathbf{B} \mathbf{R}_s \mathbf{B}^H$ and $\mathbf{r}_{bd} = \mathbf{B} \mathbf{R}_s$ and $\mathbf{r}_{db} = \mathbf{r}_{bd}^H$. The problem considered in [25] is one of estimating d given \mathbf{b} . Observe that d is a scalar. Accordingly, the error covariance matrix of eq (3.9) reduces to a scalar error

variance. Making the corresponding associations with the canonical coordinates representation, it follows that the error variance takes the form

$$\mathcal{E} = \sigma_d^2(1 - \mathbf{F}K^2\mathbf{F}^H) = \sigma_d^2(1 - K^2) \quad (3.10)$$

Clearly, reducing dimensionality increases the estimation error, i.e. K^2 resulting from a reduced dimension is less than K^2 for the full dimension problem. The approach of [25] seeks to minimize the increased error and reduce the dimensionality of the weight vector by rank ordering the projection of \mathbf{r}_{bd} onto the eigenspace of \mathbf{R}_b normalized by the corresponding eigenvalues and retaining the dominant projections.

We seek to extend the analysis for the STAP problem with unknown covariance matrix. For the STAP problem, the approach of [25] employs an estimated covariance matrix of the joint domain data. The analysis proceeds in an identical manner to that of the case of known covariance matrix with the important difference that the known covariance matrix is replaced by its maximum likelihood estimate (MLE). For Gaussian interference, the MLE is the sample covariance matrix. The method of [29] is similar to that of [25] in that it consists of matched filtering followed by interference removal using an estimate of the interference in the signal subspace from a projection of the data onto the orthogonal complement space. However, an important difference between the two methods lies in the construction of basis vectors for the orthogonal complement space. The method of [29] uses a recursive Gram-Schmidt method for determining the basis vectors. Dimensionality reduction is afforded by the use of a small number of basis vectors for the orthogonal complement space. Adaptive implementation in [29] employs the sample covariance matrix. Thus, for the methods of [25] and [29], the sample support requirement is $L \geq JN$ in general.

3.5 Entropy and Canonical Coordinates

We consider a partitioned $M \times 1$ vector $\mathbf{u} = [\mathbf{z}^T | \mathbf{p}^T]^T$. The covariance matrix of \mathbf{u} is given by

$$\mathbf{R}_u = E[\mathbf{u}\mathbf{u}^H] = \begin{bmatrix} \mathbf{R}_z & \mathbf{R}_{zp} \\ \mathbf{R}_{pz} & \mathbf{R}_p \end{bmatrix} \quad (3.11)$$

Assuming that \mathbf{u} follows a complex-Gaussian distribution, the PDF of \mathbf{u} is given by

$$f_U(\mathbf{u}) = \pi^{-M} |\mathbf{R}_u|^{-1} \exp(-q) \quad (3.12)$$

where $q = \mathbf{u}^H \mathbf{R}_u^{-1} \mathbf{u}$ and $|\cdot|$ denotes the determinant. The entropy of \mathbf{u} is given by

$$H(\mathbf{u}) = -E[\ln\{f_U(\mathbf{u})\}] = M \ln(\pi) + \ln|\mathbf{R}_u| + E(Q) \quad (3.13)$$

where $Q = \mathbf{u}^H \mathbf{R}_u^{-1} \mathbf{u} = \text{tr}[\mathbf{u}^H \mathbf{R}_u^{-1} \mathbf{u}] = \text{tr}[\mathbf{R}_u^{-1} \mathbf{u}\mathbf{u}^H]$. Upper case notation is used to denote a random variable whereas lowercase letters denote a particular value of the random variable. It follows that $E(Q) = E\{\text{tr}[\mathbf{R}_u^{-1} \mathbf{u}\mathbf{u}^H]\} = \text{tr}[\mathbf{I}] = M$. Thus,

$$H(\mathbf{u}) = M \ln(\pi) + \ln|\mathbf{R}_u| + M \quad (3.14)$$

The determinant of \mathbf{R}_u can be expressed in terms of that of \mathbf{R}_z and its Schur complement as $|\mathbf{R}_u| = |\mathbf{R}_z| |(\mathbf{R}_p - \mathbf{R}_{zp}^H \mathbf{R}_z^{-1} \mathbf{R}_{zp})|$. Using the canonical correlation representation, the determinant can be rewritten as $R_u = |\mathbf{R}_z| |(\mathbf{I} - \mathbf{F} \mathbf{K}^2 \mathbf{F}^H)| = |\mathbf{R}_z| |\mathbf{I} - \mathbf{K}^2|$. Thus, the entropy may be expressed in terms of canonical coordinates as

$$H(\mathbf{u}) = M \ln(\pi) + \ln[|\mathbf{R}_z|] + \sum_{i=1}^{\min(l, M-l)} \ln(1 - k_i^2) + M \quad (3.15)$$

where k_i denote the diagonal elements of \mathbf{K} . This last equation provides an excellent representation of entropy in terms of the canonical coordinates. Maximizing entropy is equivalent to minimizing the error covariance matrix. This also corresponds to maximizing the coherence between \mathbf{z} and \mathbf{p} .

With reference to [28], it can be shown that maximum coherence and perforce maximum entropy results when the first p lags of the clutter autocorrelation function are perfectly known. The development of entropy and canonical coordinates is useful for analyzing the work of [25], [28] and [29] in a common framework. With reference to [25, 29] only a single canonical coordinate arises since one needs to estimate the interference component in the signal subspace using a projection of the data onto the orthogonal complement space. It must be noted that the maximum value of the canonical coordinate would result when the covariance matrix is known and when full dimensionality of the noise subspace is used. The methods of [25] and [29] provide two different approaches for reducing the dimensionality of the orthogonal complement space. A criterion for determining the dimensionality employed in each algorithm can be the canonical coordinate value, K . In particular, it is desirable to have a value of K as close as possible to that obtained in the full dimension case. Once this value of K is selected, the methods of [25] and [29] employing the appropriate dimension orthogonal complement space will result in equivalent performance.

3.6 Conclusions

This chapter provides a unifying framework for three recently proposed adaptive processing methods [25, 28, 29]. Several insights using the canonical correlations method can be gained. Future work must address the issue of the canonical correlations PDF for both known and unknown covariance matrices. This will enable performance evaluation of both full and reduced dimension STAP algorithms for Gaussian interference. Extension of the canonical coordinate based analysis to STAP in spatially non-Gaussian clutter modeled by SIRPs is an important future direction as well.

Chapter 4

Adaptive Signal Processing for Non-Gaussian Clutter Backgrounds

4.1 Overview

This chapter discusses the problem of adaptive radar target detection in non-Gaussian noise backgrounds which can be modeled as a spherically invariant random process (SIRP). The estimated covariance matrix is used in an adaptive processing method for signal detection. It is shown that the resulting processor is equivalent to a generalized estimator correlator. Performance analysis is presented in terms of probability of detection versus signal-to-noise-ratio (SNR) for the case of a K-distributed SIRP. Performance comparison of our technique with that of the adaptive matched filter (AMF) of [10] is presented. The method developed in this work affords considerable performance improvement in non-Gaussian noise backgrounds.

4.2 Introduction

This investigation is motivated by the problem of adaptive signal detection in additive correlated non-Gaussian noise backgrounds. We are interested in the problem of adaptive radar target detection. Typically, the front end of a radar receiver consists of an array of antenna elements. The received signal is an electromagnetic plane wave propagating in free space, which induces a voltage at each element of the array. Usually, the received echo at the radar receiver consists of returns from one or more targets, clutter, jammer and background noise. The returns from the clutter, jammer and background noise are denoted by "noise" in this work. Relevant issues which arise as a result include target detection, estimation of unknown target amplitude, and direction of arrival

estimation. We concern ourselves with the problem of adaptive target detection in non-Gaussian noise.

Non-Gaussian statistics have been reported for scattered power from the ocean [4] and terrain [3]. This fact has been confirmed by experimental data recorded from a high resolution airborne radar operating at low grazing angles [3]. Typically most radars employ a high pulse repetition frequency (PRF). As a result, successive echoes tend to be correlated. Consequently, the joint probability density function (PDF) for the noise must include the pulse-to-pulse correlation between successive echoes. When dealing with N correlated non-Gaussian random variables, there is no unique specification of their joint PDF due to ubiquitous higher order dependencies. However, statistics which can be readily measured in practice include the first order PDF and correlation function. Hence, a model based on the first order PDF and correlation function has considerable practical merit. SIRPs are the only known class of processes which enable such a specification and have become a standard model for non-Gaussian radar clutter in recent years [6,7,15]. The SIRP model includes Gaussian clutter as a special case [6].

We address the problem of adaptive target detection in spherically invariant noise in this work. Important unknowns in this problem are the complex amplitude of the target and the noise covariance matrix. Various target models lead to different detector structures. In this work, the unknown complex amplitude of the target is modeled as a deterministic parameter. Joint maximization of the objective function (likelihood ratio) over the unknown parameters is an extremely complicated problem. A uniformly most powerful (UMP) test for this problem does not exist for the case of Gaussian noise and perforce for SIRPs. Therefore, our approach is to maximize the likelihood ratio over the unknown complex target amplitude given the maximum likelihood estimate of the covariance matrix. In this work, we show that the complex target amplitude estimate is a normalized projection of the received data vector onto the signal subspace and that the adaptive processor is equivalent to a generalized estimator correlator. It is also shown that the adaptive processor is equivalent to a matched filter compared to a data dependent threshold. Performance analysis of the adaptive processor is presented. Finally, conclusions and future research directions are pointed out.

4.3 Problem Statement

The problem of adaptive target detection in additive spherically invariant noise is formulated in the framework of a classical statistical hypothesis test between

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{y} \\ H_1 : \mathbf{x} &= a\mathbf{s} + \mathbf{y} \end{aligned} \quad (4.1)$$

where $\mathbf{x} \equiv JN \times 1$ observed complex data vector, $\mathbf{y} \equiv JN \times 1$ complex spherically invariant random vector, $\mathbf{s} \equiv$ spatio-temporal steering vector, $a \equiv$ unknown complex target amplitude, and $\Sigma \equiv$ unknown covariance matrix of \mathbf{y} .

When the covariance matrix and the target complex amplitude are known, the hypothesis test takes the form of a likelihood ratio test. However, in the presence of unknown parameters no UMP test exists. Therefore, our method consists of considering the likelihood ratio given an estimated covariance matrix. We then maximize the likelihood ratio over the unknown complex target amplitude to obtain its estimate. These estimated parameters are then used in a likelihood ratio test which forms the basis of our decision strategy. No optimality features are associated with this procedure.

Note that every SIRV y is equivalent to the product of a Gaussian random vector z and an independent non-negative random variable V , with PDF $f_V(v)$, defined to be the characteristic PDF of the SIRV. An important issue in the hypothesis testing problem of eq (4.1) is the choice of the model for the complex signal amplitude. The complex signal amplitude, a , can be modeled either as a deterministic but unknown parameter or as a random variable. In this work, a is treated as a deterministic but unknown complex signal amplitude.

4.4 Adaptive Processing Method

The hypothesis testing problem of eq (4.1) with unknown Σ and a is now considered. No optimality property of the derived test is claimed. We consider the likelihood ratio given a and Σ obtained from the PDF of the received data vector under the alternate hypotheses [14] expressed as

$$\Lambda(\mathbf{x}|a, \Sigma) = \frac{h_{2JN}(q_1)}{h_{2JN}(q_0)} \quad (4.2)$$

where

$$\begin{aligned} q_0 &= \mathbf{x}^H \Sigma^{-1} \mathbf{x} \\ q_1 &= (\mathbf{x} - a\mathbf{s})^H \Sigma^{-1} (\mathbf{x} - a\mathbf{s}) \\ h_{2JN}(w) &= \int_0^\infty v^{-2JN} \exp\left(-\frac{w}{v^2}\right) f_V(v) dv \\ H &\equiv \text{Hermitian Transpose} \end{aligned} \quad (4.3)$$

The maximization of eq (4.2) over the nuisance parameters, a and Σ , is a difficult problem in general. Again, no UMP test is available for this problem. Therefore, the approach used in this work consists of replacing the unknown parameters in eq (4.2) by their estimates.

4.4.1 Maximum Likelihood Estimate of Covariance Matrix

The unknown covariance matrix is estimated from target free training data consisting of independent identically distributed SIRVs sharing the covariance matrix of the noise in the test cell. Maximum likelihood estimation of the covariance matrix for SIRVs was first considered in [34]. The work of [34] showed that covariance matrix estimation for SIRVs can be treated in the framework of a complete-incomplete data problem and pointed out that the maximum

likelihood estimate of the covariance matrix is a weighted sample matrix. Since the problem does not permit a closed form solution, [34] uses an iterative method known as the expectation-maximization (EM) algorithm. More precisely, let \mathbf{x}_i , $i = 1, 2, \dots, K$ denote independent identically distributed training data sharing the covariance matrix of the test data vector \mathbf{x} . The work of [34] shows that the ML estimate of the covariance matrix is given by

$$\hat{\Sigma} = \frac{1}{K} \sum_{i=1}^K c_i \mathbf{x}_i \mathbf{x}_i^H \quad (4.4)$$

where $c_i = -\frac{h'_{2JN}(q_i)}{h_{2JN}(q_i)}$ with $h'_{2JN}(q) = \frac{\partial h_{2JN}(q)}{\partial q}$ and $q_i = \mathbf{x}_i^H \hat{\Sigma}^{-1} \mathbf{x}_i$. Clearly the transcendental nature of the estimate precludes obtaining a closed form solution. Consequently, [34] used the EM algorithm to obtain an iterative solution to the problem. We adopt the approach of [34] for obtaining the covariance matrix estimate in this work.

4.4.2 Maximum Likelihood Estimate of Signal Amplitude

We consider the maximization of eq (4.2) over a . Since a appears only in the numerator of eq (4.2), the maximum likelihood estimate of a is obtained from

$$\frac{\partial \Lambda(\mathbf{x}|a, \Sigma)}{\partial a} = \frac{\partial h_{2JN}(q_1)}{\partial a} = 0. \quad (4.5)$$

Note that $\frac{\partial h_{2JN}(q_1)}{\partial a} = \frac{\partial h_{2JN}(q_1)}{\partial q_1} \frac{\partial q_1}{\partial a}$. However from eq (4.3), it follows that $h_{2JN}(\cdot)$ is a monotonically decreasing function. Hence,

$$\frac{\partial h_{2JN}(q_1)}{\partial q_1} = - \int_0^\infty v^{-2JN-2} \exp(-\frac{q_1}{v^2}) f_V(v) dv < 0. \quad (4.6)$$

Hence, the desired solution is obtained by minimizing q_1 over a . We have

$$\begin{aligned} q_1 &= (\mathbf{x} - a\mathbf{s})^H \hat{\Sigma}^{-1} (\mathbf{x} - a\mathbf{s}) \\ &= \|\hat{\Sigma}^{-\frac{1}{2}} (\mathbf{x} - a\mathbf{s})\|^2 \geq 0 \end{aligned} \quad (4.7)$$

where $\|\cdot\|$ denotes the Euclidean norm. Equality in eq (4.7) holds when $a = \hat{a}$ where \hat{a} , is given by eq (4.8). The maximum likelihood estimate of a is given by

$$\hat{a} = \frac{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}}. \quad (4.8)$$

Finally, we may rewrite q_1 as

$$\hat{q}_1 = \mathbf{x}^H \hat{\Sigma}^{-1} \mathbf{x} - \frac{|\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}} = \hat{q}_0 - \frac{|\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}}. \quad (4.9)$$

where $\hat{q}_0 = \mathbf{x}^H \hat{\Sigma}^{-1} \mathbf{x}$.

4.4.3 Adaptive Processor

We derive the adaptive processor in this section and show that it is equivalent to a generalized estimator correlator. A second interpretation is that of an adaptive matched filter compared to a data dependent threshold. Using the estimates of a and Σ the likelihood ratio test takes the form

$$[\Lambda\{\mathbf{x}\}] = \left[\frac{h_{2JN}(\hat{q}_1)}{h_{2JN}(\hat{q}_0)} \right] \underset{H_0}{\overset{H_1}{>}} (T) \quad (4.10)$$

where T is a suitably chosen threshold depending on the desired false alarm probability. Let $A = |\hat{a}|^2 \mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}$. Then $h_{2JN}(\hat{q}_1)$ admits a Taylor series expansion of the form

$$h_{2JN}(\hat{q}_1) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} A^k h_{2JN}^k(\hat{q}_0) \quad (4.11)$$

where $h_{2JN}^k(q) = \frac{\partial^k h_{2JN}(q)}{\partial q^k}$. Thus, eq (4.10) takes the form

$$[\Lambda\{\mathbf{x}\}] = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} A^k \frac{h_{2JN}^k(\hat{q}_0)}{h_{2JN}(\hat{q}_0)} \quad (4.12)$$

Also, it follows that

$$f_{\mathbf{x}|V}(\mathbf{x}|v) = \pi^{-JN} |\hat{\Sigma}|^{-1} v^{-2JN} \exp(-\frac{\hat{q}_0}{v^2}) \quad (4.13)$$

$$f_{V|\mathbf{x}}(v|\mathbf{x}) = f_{\mathbf{x}|V}(\mathbf{x}|v) \frac{f_V(v)}{f_{\mathbf{x}}(\mathbf{x})} = v^{-2JN} \exp(-\frac{\hat{q}_0}{v^2}) \frac{f_V(v)}{h_{2JN}(\hat{q}_0)} \quad (4.14)$$

$$E(V^{-2k}|\mathbf{x}) = \int_0^{\infty} v^{-2k} f_{V|\mathbf{x}}(v|\mathbf{x}) dv \quad (4.15)$$

Using $f_{V|\mathbf{x}}(v|\mathbf{x})$ from eq (4.14) it follows that

$$E(V^{-2k}|\mathbf{x}) = \int_0^{\infty} v^{-2JN-2k} \exp(-\frac{\hat{q}_0}{v^2}) \frac{f_V(v)}{h_{2JN}(\hat{q}_0)} dv \quad (4.16)$$

Since $h_{2JN+2k}(q) = (-1)^k h_{2JN}^k(q)$, we have

$$E(V^{-2k}|\mathbf{x}) = (-1)^k \frac{h_{2JN}^k(\hat{q}_0)}{h_{2JN}(\hat{q}_0)} \quad (4.17)$$

This gives an interesting interpretation for eq (4.12). More precisely, the k^{th} term in the summation is the normalized product of the minimum mean squared error estimate of V^{-2k} given \mathbf{x} and the k^{th} power of A (the AMF test statistic). Thus, the test statistic of eq (4.12) is a generalized estimator-correlator. In the limit of small signal to noise ratios, powers of A go to zero much faster than A . Hence, retaining a first order approximation to the likelihood ratio produces the

adaptive locally optimum detector (ALOD) statistic as a particular case. For this specific case, the test statistic can be expressed as

$$[\Lambda\{\mathbf{x}\}] = -\frac{h'_{2JN}(\hat{q}_0)}{h_{2JN}(\hat{q}_0)}A \quad (4.18)$$

which leaves us with the interesting interpretation that the ALOD test statistic is simply the product of the Gaussian AMF test statistic and the minimum mean squared error estimate of V^{-2} given \mathbf{x} .

An alternate interpretation of the test statistic of eq (4.10) can be provided by using the monotonicity properties of $h_{2JN}(\cdot)$. More precisely, eq (4.10) can be rewritten as

$$\frac{|\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}} \underset{H_0}{\overset{H_1}{>}} T_A \quad (4.19)$$

where $T_A = \hat{q}_0 + h_{2JN}^{-1}[Th_{2JN}(\hat{q}_0)]$. This leads to the interpretation that the adaptive processor is equivalent to an adaptive matched filter compared to a data-dependent threshold. For the specific case of Gaussian noise, where $h_{2JN}(w) = \exp(-w)$, the test statistic of eq (4.10) is expressed as $\ln[\Lambda\{\mathbf{x}\}] = \ln[\frac{h_{2JN}(\hat{q}_1)}{h_{2JN}(\hat{q}_0)}] = \hat{q}_0 - \hat{q}_1 = \frac{|\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}}$ and the resulting test takes the form $\frac{|\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \hat{\Sigma}^{-1} \mathbf{s}} \underset{H_0}{\overset{H_1}{>}} T$ which is simply the AMF of [10].

4.5 Performance Analysis

Performance analysis of the test of eq (4.10) is discussed for the case of a K-distributed SIRV. The K-distribution has been proposed as a model for the amplitude statistics of ground and sea clutter observed from high resolution airborne radars operating at low grazing angles. The K-distributed envelope PDF is useful for modeling non-Gaussian radar clutter [14] and is given by

$$f_R(r) = \frac{2b}{\Gamma(\alpha)} \left(\frac{br}{2}\right)^\alpha K_{\alpha-1}(br) \quad 0 \leq r \leq \infty \quad (4.20)$$

where α is the shape parameter, b denotes the scale parameter, and $K_N(\cdot)$ is the N^{th} order modified Bessel function of the second kind. The characteristic PDF for the K-distributed SIRV and the corresponding $h_{2JN}(q)$ are

$$\begin{aligned} f_V(v) &= \frac{2b}{\Gamma(\alpha)} (bv)^{2\alpha-1} \exp(-b^2 v^2) \quad 0 \leq v \leq \infty. \\ h_{2JN}(q) &= \frac{2b^{2JN}}{\Gamma(\alpha)} (b\sqrt{q})^{\alpha-JN} K_{\alpha-JN}(2b\sqrt{q}). \end{aligned} \quad (4.21)$$

Performance evaluation is carried out using Monte-Carlo simulation. Data from the K-distributed SIRP is generated using the technique of [7]. Figure 2.2 shows the performance of the adaptive matched filter (AMF) of [10] in K-distributed SIRV. Curve 1 corresponds to the performance of the AMF operating in Gaussian noise with known covariance matrix. This is the best achievable

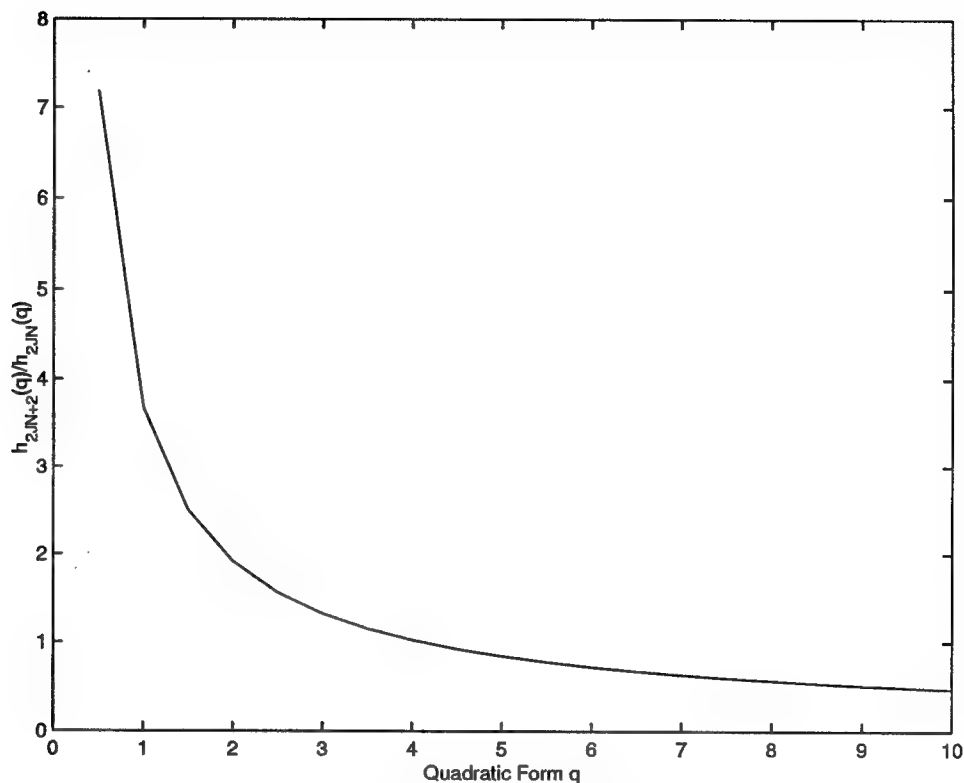


Figure 4.1: Minimum Mean Squared Error Estimate of V^{-2}

performance in Gaussian noise. Curve 2 corresponds to AMF performance in Gaussian noise using an estimated covariance matrix from training data. Relevant test parameters are reported in the figure. Curve 2 is generated using the procedure described in [10]. Curve 4 shows the AMF performance while operating in K-distributed clutter with shape parameter 0.5. AMF performance is severely degraded. This is due to the fact that the sample covariance matrix is no longer the ML estimate of the covariance matrix. The non-ergodicity of the SIRP causes a tremendous increase in training data support ($K=10JN$) to obtain performance comparable to that shown in curve 2. This is shown in curve 3.

Figure 4.1 shows a plot of c_i as a function of q . Observe that c_i , which denotes the MMSE estimate of V_i^{-2} given \mathbf{x}_i , is the ratio of two monotonically decreasing functions of q and hence is monotonically decreasing. This factor corrects for the non-ergodicity of the SIRP. Figure 4.2 shows the performance of the adaptive processing method of eq (4.10) for the case of a K-distributed SIRV using the ML estimate of the covariance matrix. Relevant test parameters are provided on the plot caption. Significant performance improvement results when

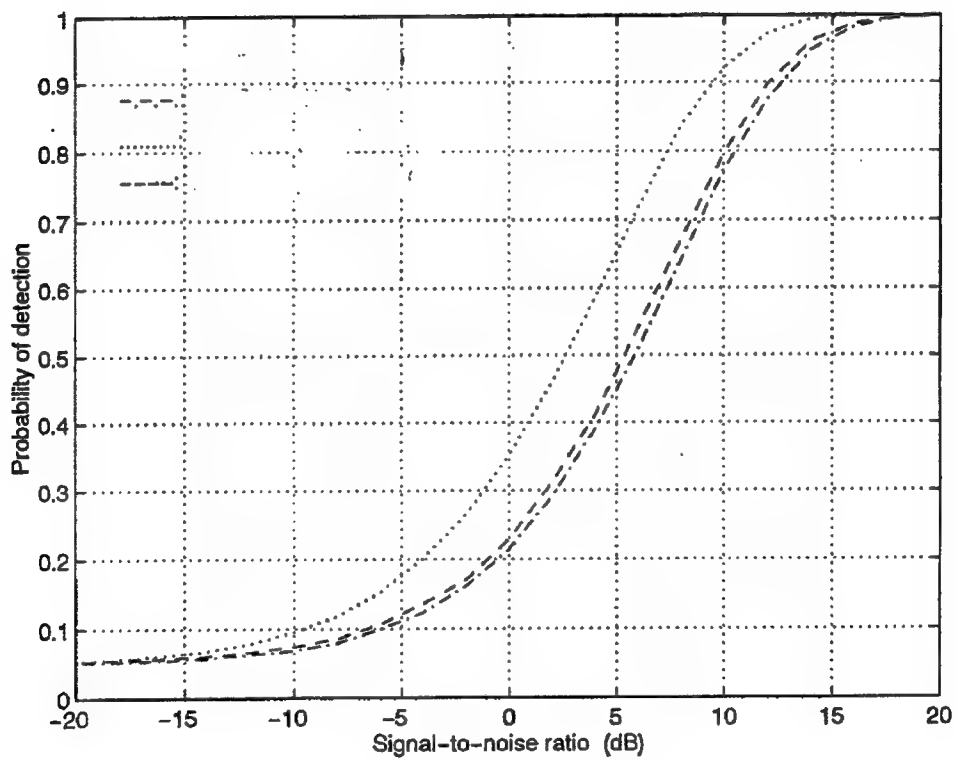


Figure 4.2: Probability of Detection versus SINR $K=10$, $J=1$, $N=5$ $P_{fa} = 0.01$: 1=Adaptive Processor of eq(4.10) in K-distributed Noise $\alpha = 1$, $b = 1$, 2=Matched Filter in Gaussian Noise, 3=AMF in Gaussian Noise

the ML estimate of the covariance matrix is used. The performance approaches that of the AMF operating in Gaussian noise. However, there is still some performance penalty (approximately 1 dB SNR for $P_d = 0.5$) compared to the AMF operating in Gaussian noise. This is due to the non-Gaussian statistics of the observed data.

4.6 Conclusions

In this work, we derived the adaptive processor for target detection in non-Gaussian noise backgrounds. The adaptive processor is shown to be a generalized estimator-correlator. We also showed that the performance of the AMF in non-Gaussian noise is severely degraded. Our method demonstrates considerable performance improvement in non-Gaussian noise backgrounds and shows that using 2JN training data vectors, performance to within 1 dB of the AMF for Gaussian noise can be obtained. Even the 2JN requirement can be quite stringent in several practical applications. This can be alleviated by using the model based parametric method of [28], and the PCI method of [27].

Chapter 5

Adaptive Signal Processing in Structured Interference Backgrounds

5.1 Overview

Important issues in space-time adaptive processing are the reduction of computational complexity and training data support for covariance matrix estimation. Sample covariance matrix based methods used in [8]- [13] have a computational cost of $O(M^3)$, where $M=JN$ is the spatio-temporal product. Additionally $3dB$ SNR performance requires $K = 2JN$ independent, identically distributed training data vectors for covariance matrix estimation. However in many instances, a priori information about the structure of the interference background may be available. This information can be used to reduce the training data support and computational complexity of the adaptive processing method. A commonly encountered disturbance scenario in radar and sonar signal processing is strong low-rank interference plus white noise. For this scenario, the principal components inverse (PCI) developed by Tufts [26, 27] provides excellent performance while reducing the training data support and the computational complexity. In this work, the PCI method of [26, 27] is applied to non-Gaussian interference phenomena modeled by SIRPs. It is seen that the resulting algorithm performance degrades significantly when an estimated subspace is used. This is due to increased subspace perturbation caused by the non-ergodicity of the SIRP.

5.2 Introduction

This work is motivated by the problem of adaptive target detection in non-homogeneous and non-Gaussian interference scenarios when a priori information concerning the structure of the interference is available. A typical radar receiver

front end consists of an array of antenna elements. The measured data at the array arises from an electromagnetic plane wave impinging on the array manifold. The plane wave induces a voltage at each element of the array, which gives rise to the measured data. Given the measured data, the problem of interest is to adaptively detect targets against a background of interference consisting of clutter, jammer and background noise with unknown covariance matrix. This problem for the case of Gaussian interference has received considerable attention [8]- [13]. These efforts are based on the inversion of a sample covariance matrix which is the maximum likelihood estimate for the case of Gaussian interference. Sample matrix inversion has a computational cost of $O(M^3)$ where $M = JN$ is the spatio-temporal product. Furthermore, positive definiteness of the estimated covariance matrix requires $K \geq JN$ independent, identically distributed, representative training data vectors in forming the sample matrix. The work of [8] also shows that to obtain performance within 3 dB of the optimal SNR, $K \approx 2JN$ training data vectors are needed. The computational cost and training data support requirements become onerous in the limit of large JN . Consequently, important issues in present-day STAP processing include reduction of computational cost and training data support.

In many instances, a priori information concerning the interference structure may be available. For example, several radar clutter measurements tend to support an interference model of the form of a strong low rank component plus white background noise. In such instances, significant reduction in training data support and computational cost is possible. The principal components inverse (PCI) method developed by Tufts [35] provides a useful approach for target detection in these scenarios. The main result of [35] is the derivation of the PDF for the sum of the squares of the singular values when the signal subspace and orthogonal complement space are estimated using finite data. Application of the PCI method for adaptive target detection in Gaussian interference can be found in [26,27]. In this work, the PCI method is applied to non-Gaussian interference described by spherically invariant random processes (SIRP). It is seen that PCI performance in SIRP interference degrades severely compared to performance in Gaussian interference. This is due to increased subspace perturbation caused by the non-ergodicity of the SIRP.

This work was carried out in collaboration with Prof. Donald Tufts and Mr. Brian Freburger of the Electrical Engineering Department of the University of Rhode Island. This collaboration was a result of the AFOSR/DSTO meeting at Victor Harbour, Australia, June 1997. The main results of this chapter were presented in an invited paper at the conference on information sciences and systems (CISS), Princeton, NJ, March 1998. This paper is attached to the report in Appendix A. The author is most grateful to Prof. Tufts and Mr. Freburger for the many stimulating discussions on the PCI method and its application to non-Gaussian SIRPs. Since details of the research carried out are contained in the paper, the following sections will provide a brief outline of the work.

5.3 Problem Statement

We consider the problem of adaptive target detection of a signal known to within a multiplicative constant of a steering vector in interference with unknown spectral characteristics. The hypothesis test used is

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{n} \\ H_1 : \mathbf{x} &= \alpha \mathbf{s} + \mathbf{n} \end{aligned} \quad (5.1)$$

where \mathbf{x} denotes the observed complex data vector, \mathbf{s} is a known steering vector, α is the unknown complex target amplitude, and \mathbf{n} is a zero mean complex valued interference vector whose covariance matrix $\mathbf{R} = E[\mathbf{n}\mathbf{n}^H]$ is typically unknown. For this analysis, it is assumed that $\mathbf{R} = \mathbf{Q} + \sigma^2 \mathbf{I}$ where \mathbf{Q} is a low rank matrix with large eigenvalues compared to σ^2 . Since \mathbf{R} is typically unknown, its sample matrix estimate formed from target-free representative training data is used. The sample matrix estimate is a maximum likelihood estimate of the covariance matrix for Gaussian interference [8] when no restrictions besides the positive definite Hermitian condition are imposed on the structure of the covariance matrix. An $M \times M$ sample matrix requires at least $K = M$ independent identically distributed, target-free training data to guarantee its positive definiteness [36]. The work of [8] shows that to obtain an SNR within 3 dB of the optimal performance using an estimated covariance matrix, $K = 2M$ training data vectors are needed. Furthermore, the sample matrix inversion technique of [8]- [13] has a computational complexity of $O(M^3)$. These requirements render the sample matrix inversion approach inefficient for structured covariance estimation.

5.4 The PCI Method

The principal components inverse (PCI) method proposed in [27] and the references therein avoids formation and explicit inversion of the sample covariance matrix. We shall see how this is accomplished in the following. Recall that the optimal weight vector in the Reed-Mallet-Brennan (RMB) beamformer of [8] is given by $\mathbf{w} = \mathbf{R}^{-1} \mathbf{v}$ and the test statistic of the optimal beamformer is simply $T = |\mathbf{w}^H \mathbf{x}|^2 = |\mathbf{s}^H \mathbf{R}^{-1} \mathbf{x}|^2$. Since \mathbf{R} is a Hermitian positive definite matrix, it admits a representation of the form

$$\mathbf{R} = \mathbf{W} \mathbf{D} \mathbf{W}^H = \sum_{i=1}^M \lambda_i \mathbf{w}_i \mathbf{w}_i^H \quad (5.2)$$

where λ_i and \mathbf{w}_i denote the eigenvalues and eigenvectors of \mathbf{R} . If \mathbf{R} is made up of a small number of low rank interferers plus white noise, the representation for $\mathbf{R} = \mathbf{B} + \sigma^2 \mathbf{I}$ takes on the form $\mathbf{R} = \sum_{i=1}^k (\lambda_i + \sigma^2) \mathbf{w}_i \mathbf{w}_i^H + \sum_{i=k+1}^M \sigma^2 \mathbf{w}_i \mathbf{w}_i^H$, where k is the rank of \mathbf{B} . Accordingly, \mathbf{R}^{-1} takes on the form

$$\begin{aligned} \mathbf{R}^{-1} &= \sum_{i=1}^M (\lambda_i + \sigma^2)^{-1} \mathbf{w}_i \mathbf{w}_i^H \\ &= \frac{1}{\sigma^2} \sum_{i=1}^M \left(\frac{\lambda_i}{\sigma^2} + 1 \right)^{-1} \mathbf{w}_i \mathbf{w}_i^H \\ &= \frac{1}{\sigma^2} (\mathbf{I} - \mathbf{P}) \end{aligned} \quad (5.3)$$

where

$$\mathbf{P} = \sum_{i=1}^M \frac{\lambda_i}{\sigma^2} \left(\frac{\lambda_i}{\sigma^2} + 1 \right)^{-1} \mathbf{w}_i \mathbf{w}_i^H \quad (5.4)$$

When $\lambda_i \gg \sigma^2$, the matrix \mathbf{P} can be approximated as

$$\mathbf{P} \approx \sum_{i=1}^k \mathbf{w}_i \mathbf{w}_i^H \quad (5.5)$$

where k is the rank of \mathbf{B} . Note that \mathbf{w}_i , $i = 1, 2, \dots, k$, are the eigenvectors corresponding to the dominant eigenvalues of \mathbf{R} . The test statistic of the RMB beamformer can be expressed as

$$T = |\mathbf{s}^H \frac{1}{\sigma^2} (\mathbf{I} - \mathbf{P}) \mathbf{x}|^2 \quad (5.6)$$

The test statistic can then be interpreted as the energy of the matched filter output where the projection of the received data onto the orthogonal complement space of the dominant interference space is match filtered with the steering vector [27].

In practice, \mathbf{R} and \mathbf{w}_i are unknown and must be estimated from representative training data. However, estimation of \mathbf{R} is not required to estimate \mathbf{w}_i . Estimation of \mathbf{w}_i can be carried out directly using a singular value decomposition (SVD) of the data matrix of target-free training data [26]. Significant computational savings and reduction of training data support result when k is small. A key issue in this process is the determination of k . Tufts showed in [35] that the sum of the squares of the singular values of the target-free training data matrix follows a chi-square distribution. Based on this result, it is possible to determine k in a straightforward manner [35]. However since determination of k is scenario dependent, it must be carried out on a case-by-case basis. The test statistic of the form of eq (5.6) is not CFAR since the SNR produced by this test incurs a dependence on the true covariance matrix of the interference scenario. In a recent effort, Reed and Gau derived the CFAR version of the PCI in [37]. However, this work assumes that a perfect estimate of \mathbf{w}_i , $i = 1, 2, \dots, k$ and hence, of \mathbf{P} is available. Therefore, derivation of a PCI-based CFAR test statistic for estimated subspaces is still an open problem.

Performance of the PCI method is characterized in two ways. The first method is based on subspace perturbation. Since, the dominant eigenvectors are estimated using finite training data support, the estimated dominant interference subspace is viewed as a perturbation of the true dominant subspace [38]. Perturbation of the true subspace is reflected by the probability density function (PDF) of the sum of the squares of the estimated singular values. A second way of characterizing PCI performance is via a subspace swap. In practice, the dominant interference subspace is estimated from finite training data. Consequently, on any given realization, it is likely that a linear combination of the orthogonal complement subspace can resolve more energy than an eigenvector in the dominant interference space. In such a case, a swap is said to have taken

place. Analysis of the swap probability was first carried out in [39]. A lower bound on the swap probability was derived in [40].

5.5 PCI for Non-Gaussian Interference

The PCI method was applied to strong low-rank non-Gaussian interference modeled by a K-distributed SIRP. The PCI test statistic of eq (5.6) is formed and then compared to a suitably determined threshold. Since the PDF of the test statistic in eq (5.6) for estimated \mathbf{P} is unknown in closed form, it is not possible to calculate the probability of detection and probability of false alarm in closed form or via integral expressions. Hence, performance analysis is carried out using Monte Carlo simulations. Details of the simulation are described in [41] which is attached as Appendix A of this report. For the case of a known covariance matrix and hence known \mathbf{P} , PCI performance in non-Gaussian interference is identical to the performance in Gaussian interference. However, for estimated \mathbf{P} , PCI performance in non-Gaussian interference degrades significantly compared to the performance in Gaussian interference particularly for small sample sizes and low SNR. This is due to the non-ergodicity of the SIRP. Details of the approach, representative examples, and discussion of the results are available in [41].

5.6 Conclusion

The PCI method provides a useful technique for overcoming limitations of computational complexity and training data support for adaptive processing in interference backgrounds made up of a low-rank component plus white noise. However, when applied to non-Gaussian SIRP interference scenarios significant performance degradation is encountered. Techniques for overcoming this degradation can be developed by viewing the PCI method for non-Gaussian SIRPs in the framework of an incomplete data problem. Performance restoration to within 1 dB of the corresponding PCI performance for Gaussian interference is possible. This is an area of future investigation.

Chapter 6

Conclusions

This effort addressed the problem of adaptive target detection in an interference background made up of clutter, jammer, and background noise. Most classical STAP methods are based on sample covariance matrix inversion. These methods require considerable amounts of training data for covariance matrix estimation and have a computational cost of $O(M^3)$ where $M = JN$ is the spatio-temporal product. The performance of these methods degrades severely in non-homogeneous/ non-Gaussian interference environments.

This study investigated techniques for reducing the training data support and computational cost of STAP algorithms for non-Gaussian interference scenarios. The key feature of this study was to show that model based parametric STAP methods offer the potential for considerable performance improvement by reducing training data support. Another important result is the performance improvement in non-Gaussian interference scenarios.

Important issues in the context of transitioning the results of this study into ongoing programs at U.S. AFRL such as MCARM and AASP include:

1. The application of the techniques developed herein to real data.
2. Account for the effects of platform velocity, crab angle, mutual coupling between the antenna array elements, and internal clutter motion.
3. Carry out detailed investigations of the CFAR properties of the PAMF, PCI and other reduced-rank STAP methods.
4. Analyze the robustness features due to model mismatch.

These issues will be the subject of investigation in future studies.

Bibliography

- *[1] A. Jaffer, M. Baker, W. Balance, and J. Staub, "Adaptive space-time processing techniques for airborne radar," Tech. Rep. RL-TR-91-162, Rome Laboratory, July 1991.
- [2] J. Ward, "Space-time adaptive processing for airborne radar," Tech. Rep. Technical Report 1015, MIT Lincoln Laboratory, December 1994.
- [3] J. Jao, "Amplitude distribution of composite terrain radar clutter and the K-distribution," *IEEE Trans. on Antennas and Propagation*, vol. AP-32, pp. 1049-1062, 1984.
- [4] E. Jakeman and P. Pusey, "A model for non-Rayleigh sea echo," *IEEE Trans. on Antennas and Propagation*, vol. AP-24, pp. 806-814, 1976.
- [5] E. Conte and M. Longo, "Characterization of radar clutter as a spherically invariant random process," *IEE Proc.F, Commun., Radar, & Signal Process.*, vol. 134, (2), pp. 191-197, 1987.
- [6] M. Rangaswamy, D. D. Weiner, and A. Ozturk, "Non-Gaussian random vector identification using spherically invariant random processes," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-29, pp. 111-124, 1993.
- [7] M. Rangaswamy, D. D. Weiner, and A. Ozturk, "Computer generation of correlated non-Gaussian radar clutter," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-31, pp. 106-116, 1995.
- [8] I. Reed, J. Mallett, and L. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-10, pp. 853-863, 1974.
- [9] W. Chen and I. Reed, "A new CFAR detection test for radar," *Digital Signal Processing*, vol. 1, pp. 198-214, 1991.
- [10] F. Robey, D. Fuhrmann, E. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-28, pp. 208-216, 1992.

*RL-TR-91-162 is Distribution Limited to US Gov't Agencies and their contractors.

- [11] L. L. Scharf and B. Friedlander, "Matched Subspace Detectors," *IEEE Trans. on Signal Processing*, vol. SP-42, pp. 2146-2157, 1994.
- [12] L. L. Scharf and T. L. McWhorter, "Adaptive Matched Subspace Detector and Adaptive Coherence Estimators," in *Proceedings of the 30th Asilomar Conference on Signals, Systems, and Computers*, (Pacific Grove, CA), 1996.
- [13] L. L. Scharf, T. L. McWhorter, and L. J. Griffiths, "Adaptive Coherence Estimation for Radar Signal Processing," in *Proceedings of the 30th Asilomar Conference on Signals, Systems, and Computers*, (Pacific Grove, CA), 1996.
- [14] M. Rangaswamy, J. H. Michels, and D. D. Weiner, "Multichannel detection for correlated non-Gaussian random processes based on innovations," *IEEE Trans. on Signal Processing*, vol. SP-43, pp. 1915-1922, 1995.
- [15] K. Sangston and K. Gerlach, "Coherent detection of targets in a non-Gaussian background," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-30, pp. 330-340, 1994.
- [16] S. Kay, *Modern Spectral Estimation-Theory and Application*. New Jersey: Prentice Hall, 1988.
- [17] E. Conte, M. Lops, and G. Ricci, "Adaptive matched filter detection in spherically invariant noise," *IEEE Signal Processing Letters*, vol. 3, No. 8, pp. 248-250, 1996.
- [18] C. Therrien, "On the relation between triangular matrix decomposition and linear prediction," *Proceedings of the IEEE*, vol. 71, no. 12, pp. 1459-1460, 1983.
- [19] J. Michels, "Covariance Matrix Estimator Performance in Non-Gaussian Clutter Processes," in *Proceedings of the IEEE National Radar Conference*, (Syracuse, NY), 1997.
- [20] M. Rangaswamy and J. H. Michels, "Adaptive Processing in Non-Gaussian Noise Backgrounds," in *Proceedings of the Ninth IEEE Workshop on Statistical Signal and Array Processing*, (Portland, OR), 1998.
- [21] O. Strand, "Multichannel complex maximum entropy (auto-regressive) spectral analysis," *IEEE Trans. on Automatic Control*, vol. AC-22, pp. 634-640, 1977.
- [22] A. Nuttall, "Multivariate Linear Predictive Spectral Analysis Employing Weighted Forward and Backward Averaging," Tech. Rep. TR-5501, Naval Underwater Systems Center, New London, CT, October 1976.
- [23] S. Applebaum, "Adaptive arrays," tech. rep., Syracuse University Research Corporation, 1964.

- [24] L. E. Brennan and I. S. Reed, "Theory of adaptive radar," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-9, pp. 237-252, 1973.
- [25] J. Goldstein and I. Reed, "Theory of Partially Adaptive Radar," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-33, no.4, pp. 1309-1325, 1997.
- [26] I. KIRSTEINS and D. TUFTS, "Adaptive detection using a low rank approximation to a data matrix," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-30, pp. 55-67, 1994.
- [27] D. Tufts and B. Freburger, "Rapidly Adaptive Signal Detection Using the Principal Component Inverse (PCI) Method," *Signal Processing*, 1998.
- [28] M. Rangaswamy and J. H. Michels, "A Parametric Detection Algorithm for Space-Time Adaptive Processing in Non-Gaussian Radar Clutter Backgrounds," *Signal Processing*, 1998.
- [29] D. Pados, J. Michels, T. Tsao, and M. Wicks, "Joint Domain Space-Time Adaptive Processing With Small Training Data Sets," in *Proceedings of the IEEE National Radar Conference*, (Dallas, TX), 1998.
- [30] J. Roman, "Adaptive Sidelobe Cancelling Using Complex Valued Canonical Variables." AFRL-SN-RS-TR-1998-46, March 1998.
- [31] L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley Publishing Company, 1990.
- [32] L. L. Scharf and J. K. Thomas, "Wiener Filters in Canonical Coordinates for Transform Coding, Filtering, and Quantizing," *IEEE Trans. on Signal Processing*, vol. SP-46, pp. 647-654, 1998.
- [33] H. Hotelling, "Relations Between Two Sets of Variates," *Biometrika*, vol. 28, pp. 321-377, 1936.
- [34] R. J. Little and D. B. Rubin, *Statistical Analysis With Missing Data*. New York: John Wiley & Sons, 1987.
- [35] D. Tufts and A. A. Shah, "Determination of a Signal Subspace From Short Data Records," *IEEE Trans. on Signal Processing*, vol. 42, pp. 2531-2535, 1994.
- [36] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of structured covariance matrices," *Proceedings of the IEEE*, vol. 70, No.9, pp. 963-974, 1982.
- [37] Y.-L. Gau and I. S. Reed, "An improved reduced-rank CFAR space-time adaptive radar detection algorithm," *IEEE Trans. on Signal Processing*, vol. SP-46, pp. 2139-2146, 1998.

- [38] R. J. Vaccaro, "A second order perturbation expansion for the SVD," *SIAM Journal on Matrix Analysis Applications*, vol. 15,(2), pp. 661-671, 1994.
- [39] D. W. Tufts, A. C. Kot, and R. J. Vaccaro, "The analysis of threshold behavior of SVD based algorithms," in *Proceedings of the 21st Asilomar Conference on Signals, Systems, and Computers*, (Pacific Grove, CA), pp. 550-554, 1987.
- [40] J. K. Thomas, L. L. Scharf, and D. W. Tufts, "The probability of a subspace swap in the SVD," *IEEE Trans. on Signal Processing*, vol. 43, pp. 730-736, 1995.
- [41] B. E. Freburger, D. W. Tufts, and M. Rangaswamy, "Signal Detection Performance of Principal Components Inverse (PCI) For Strong Low Rank Non-Gaussian Interference," in *Proceedings of thirtysecond Conference on Information Sciences and Systems*, (Princeton, NJ), 1998.

Appendix A: PCI Performance in Non-Gaussian Interference Backgrounds

SIGNAL DETECTION PERFORMANCE OF PRINCIPAL COMPONENTS INVERSE (PCI) FOR STRONG, LOW RANK, NON-GAUSSIAN INTERFERENCE

*Brian E. Freburger and Dr. Donald W. Tufts **

Dr. Muralidhar Rangaswamy †

Department of Electrical and Computer Engineering
University of Rhode Island
Kelley Hall
Kingston, RI 02881

ARCON Corporation
260 Bear Hill Rd.
Waltham, MA 02154

ABSTRACT

This paper examines the performance of the Principal Components Inverse (PCI) [12, 4, 6, 5] method of adaptive detection in the presence of strong low rank non-Gaussian interference. The non-Gaussian interference is generated using a spherically invariant random vector model [8, 9, 10]. With strong interference, near optimal interference suppression can be achieved by null projecting the data away from the interference subspace. When using the known projection, performance is identical for Gaussian and non-Gaussian interference. However, the performance for non-Gaussian noise degrades much faster than the Gaussian performance when the projection is estimated. The effect is most noticeable for small sample sizes and low signal levels.

1. PROBLEM STATEMENT

Consider the problem of detecting a signal of known shape but unknown complex amplitude embedded in noise. The hypothesis test is

$$H_0 : X = N \quad (1)$$

$$H_1 : X = \alpha S + N \quad (2)$$

where X is a vector of complex valued data, S is a known signal shape vector and N is a zero mean complex random vector with covariance matrix $R = E(NN^H)$. The complex signal amplitude α is assumed to be deterministic but unknown. Also, assume that R is the combination of a strong low rank process and white background noise such that $R = Q + \sigma^2 I$ where Q is low rank with eigenvalues large compared to σ^2 . This scenario is typical of problems in radar and active sonar where X represents a spatial/temporal sampling.

In the case of unknown R , it is typically assumed that there exists a set of secondary data which can be used to estimate the covariance. The method of Principal Components Inverse (PCI) is often employed for adaptive detection in this scenario when the interference model is Gaussian. However, often the interference is not Gaussian, and

the performance of applying this detection method needs to be characterized for non-Gaussian interference.

2. NON-GAUSSIAN NOISE MODEL

In the above, the white background noise associated with the $\sigma^2 I$ portion of the covariance will be assumed to always be zero mean Gaussian. The portion associated with Q may often be from a non-Gaussian process. Non-Gaussian noise/disturbance phenomena have been reported in applications like radar, sonar, digital communications, and radio astronomy. Most of the analysis in these applications is based on the assumption of independent identically distributed random variables. However, this assumption fails in many instances. For example, coherent processing in radars operating at a high pulse repetition frequency (PRF) results in successive radar returns (pulses) being highly correlated. Consequently, it is necessary to describe the joint probability density of N correlated non-Gaussian random variables. It must be noted that unlike the Gaussian case, the PDF of N correlated non-Gaussian random variables is not unique. This is due to the large number of ubiquitous higher order dependencies. However, only the first order PDF and correlation function can be readily measured in practice. Thus, there is considerable practical merit to a mechanism which affords the specification of the joint PDF of N correlated non-Gaussian random variables based on the first order PDF and a covariance matrix. SIRPs are the only known class of processes which permit such a specification. SIRPs are the multivariate extension of compound models which have been found to satisfactorily account for the first order PDF of experimentally measured terrain and sea clutter which exhibit spiky or impulsive behavior.

We present definitions and relevant mathematical preliminaries for complex-SIRVs and complex-SIRPs in this section. A zero mean random vector $\tilde{Y} = Y_c + jY_s$, where $Y_c = [Y_{c1}, Y_{c2}, \dots, Y_{cN}]^T$ and $Y_s = [Y_{s1}, Y_{s2}, \dots, Y_{sN}]^T$ denote the vectors of the in-phase and out-of-phase quadrature components, is a complex-SIRV if its PDF has the form

$$f_{\tilde{Y}}(\tilde{y}) = (\pi)^{-N} |\tilde{\Sigma}|^{-1} h_{2N}(p) \quad (3)$$

where $p = \tilde{y}^H \tilde{\Sigma}^{-1} \tilde{y}$, $\tilde{\Sigma}$ is a non-negative definite Hermitian matrix, and $h_{2N}(\cdot)$ is a positive, real valued, monoton-

Dr. Tufts and Mr. Freburger were supported by ONR under U.S. Navy contract no: N000149610938

Dr. Rangaswamy was supported by the AFOSR under U.S. Air Force Research Laboratory contract no: F-30602-97-C-0050

ically decreasing function for all N . If every random vector obtained by sampling a complex random process, $\tilde{y}(t)$, is a complex-SIRV, regardless of the sampling instants or the number of samples, then the process $\tilde{y}(t)$ is defined to be a complex SIRP.

Yao in [15] derived a representation theorem for real SIRVs. The representation theorem extends to complex-SIRVs readily and is stated as follows.

The random vector \tilde{Y} is a complex-SIRV if and only if it is equivalent to the product of a complex-Gaussian random vector \tilde{Z} and a non-negative random variable V with PDF $f_V(v)$, defined to be the characteristic PDF of the complex-SIRV.

Consequently,

$$\begin{aligned}\tilde{Y} &= \tilde{Z}V \\ f_{\tilde{Y}}(\tilde{y}) &= \int_0^\infty f_{\tilde{Y}|V}(\tilde{y}|v) f_V(v) dv \\ f_{\tilde{Y}|V}(\tilde{y}|v) &= \pi^{-N} |\tilde{\Sigma}|^{-1} v^{-2N} \exp(-\frac{p}{v^2}) \\ h_{2N}(p) &= \int_0^\infty v^{-2N} \exp(-\frac{p}{v^2}) f_V(v) dv.\end{aligned}\quad (4)$$

It is assumed without loss of generality that $E(V^2) = 1$ so that the covariance matrix of the complex-SIRV is equal to that of the complex-Gaussian random vector. Due to the assumption $E(V^2) = 1$, the covariance matrix of the complex-SIRV is $\tilde{\Sigma}$.

The representation theorem and the assumption that $E(V^2) = 1$, give rise to the following necessary and sufficient conditions for representing \tilde{Y} as a complex-SIRV

$$\begin{aligned}E\{Y_c\} &= E\{Y_s\} = 0 \\ \Sigma_{cc} &= \Sigma_{ss} \\ \Sigma_{cs} &= -\Sigma_{sc}.\end{aligned}\quad (5)$$

where

$$\begin{aligned}\Sigma_{cc} &= E\{Y_c Y_c^T\} & \Sigma_{ss} &= E\{Y_s Y_s^T\} \\ \Sigma_{cs} &= E\{Y_c Y_s^T\} & \Sigma_{sc} &= E\{Y_s Y_c^T\}.\end{aligned}\quad (6)$$

Under these conditions, it follows that

$$\tilde{\Sigma} = 2[\Sigma_{cc} + j\Sigma_{sc}]. \quad (7)$$

Several attractive properties of complex-Gaussian random vectors generalize to complex-SIRVs as a consequence of the representation theorem. Complex-SIRVs satisfying the conditions of eq (5) are also called SIRVs of the circular class [7]. An important property of complex-SIRVs of the circular class is the linearity of estimators in minimum mean square error estimation (MMSE) problems [7].

3. PRINCIPAL COMPONENTS INVERSE (PCI)

3.1. Optimum Detector

When the signal is embedded in noise which has a multivariate Gaussian probability distribution with zero mean vector and known covariance matrix R , then the maximum signal to noise ratio test statistic [2] is a function of the magnitude of the output of a noise whitening matched filter [1]

$$|S^H R^{-1} X| \quad (8)$$

3.2. Gaussian Based Development

Assume the scenario as given in the problem statement of section 1.. The noise whitened matched filter can be written in two equivalent forms [3].

$$|S^H R^{-1} X| = \frac{1}{\sigma^2} |S^H (I - P_w) X| \quad (9)$$

where P_w is a weighted projection matrix given by

$$P_w = \sum_{k=1}^n \frac{\lambda_k}{\lambda_k + \sigma^2} Q_k Q_k^H \quad (10)$$

where Q_k is the k^{th} singular vector of Q and λ_k is the associated singular value.

The Sample Matrix Inverse (SMI) method estimates R with $\hat{R} = \frac{1}{K} \sum_{k=1}^K X_k X_k^H$ and uses this in the test given by

$$|S^H \hat{R}^{-1} X| \quad (11)$$

The PCI method takes the approach of estimating P_w and using

$$|S^H (I - \hat{P}_w) X| = |S^H (X - \hat{P}_w X)| \quad (12)$$

If we further assume that the low rank portion of the noise is much stronger than the part of the noise that is white then for a Q matrix of rank r we have

$$\lambda_k \gg \sigma^2 \text{ for } 1 \leq k \leq r \text{ and } \lambda_k \approx 0 \text{ for } k > r \quad (13)$$

and the weighted projection can be well approximated by a standard projection matrix

$$P_w \approx P = \sum_{k=1}^r Q_k Q_k^H \quad (14)$$

The PCI method can then be summarized as estimate the singular vectors, \hat{Q}_k , and use

$$|S^H (I - \sum_{k=1}^r \hat{Q}_k \hat{Q}_k^H) X| \quad (15)$$

If a matrix of the data snapshots is formed

$$X = [X_1 \ X_2 \ \dots \ X_K] \quad (16)$$

then the left singular vectors of \hat{R} and the matrix X are the same and the singular vectors, \hat{Q}_k , can be calculated directly from the data matrix X .

3.3. Rank Selection Strategy

One of the goals of the PCI method is that under conditions of no strong interference, the waveforms should pass through the PCI processing unchanged. To this end, a threshold value, T , is set which is compared to the partial sums of the singular values of the data matrix. That is, assuming the singular values to be in descending order, $\sigma_1 > \sigma_2 > \dots > \sigma_N$ we seek the smallest S, S_{min} for which

$$\sum_{k=0}^S \sigma_{N-k}^2 > T \quad (17)$$

The rank is then chosen as $r = N - S_{min}$ where N is the row dimension of the space-time data matrix and T is the threshold value. When the value of $S_{min} = N$ then the rank is chosen as 0 and the PCI processing is turned off so that no interference suppression is performed. The threshold is chosen such that low level noise and the max expected target level (which is usually range dependent) will be within the threshold T . This choice satisfies the previously stated goal of passing the antenna waveforms unchanged in the absence of strong jamming.

3.4. PCI Performance

The performance of the PCI method can be characterized in two ways. The first is called a subspace swap [13, 11]. In the case where we have the notion of a correct interference rank, separate the eigenvectors of the covariance matrix $\mathbf{R} = \mathbf{U}\mathbf{A}\mathbf{U}^H$ into interference eigenvectors and orthogonal eigenvectors

$$\mathbf{U} = [\mathbf{U}_s \mid \mathbf{U}_o] \quad (18)$$

$$= [U_1 \cdots U_r \mid U_{r+1} \cdots U_N] \quad (19)$$

where U_k is the eigenvector associated with the k^{th} largest eigenvalue. For a given realization of the data matrix, one expects the energy to be greater in the direction of the eigenvectors of the interference subspace than in the direction of any linear combination of the eigenvectors of the orthogonal subspace. Since any unit length vector in the orthogonal subspace can be expanded as $U_o a$ we have,

$$\|U_i U_i^H \mathbf{X}\|_F > \sup_{\|a\|=1} \|U_o a a^H U_o^H \mathbf{X}\|_F \text{ for } i = 1 \cdots r \quad (20)$$

However, because of fluctuations in finite observations of data, it can happen that a linear combination of the eigenvectors in the orthogonal space resolves more energy than a eigenvector in the signal space. This is called a subspace swap and is associated with a rapid degradation in the performance of SVD based algorithms [13].

When a subspace swap does not occur, the estimated subspace can be viewed as a perturbation of the true subspace. Let the matrix of background noise realizations be viewed as a perturbation such that the noise matrix $\mathbf{N} = [\mathbf{N}_1 \ \mathbf{N}_2 \ \cdots \ \mathbf{N}_K]$ can be written as $\mathbf{N} = \mathbf{Z} + \mathbf{W}$ where \mathbf{Z} is the low rank interference and \mathbf{W} is the background Gaussian noise. Letting the SVD of \mathbf{Z} be

$$\mathbf{Z} = [\mathbf{U}_s \ \mathbf{U}_o] \begin{bmatrix} \Sigma_s & 0 \\ 0 & \Sigma_o \end{bmatrix} [\mathbf{V}_s \ \mathbf{V}_o] \quad (21)$$

the PCI projection matrix can be approximated by [14]

$$(\mathbf{I} - \tilde{\mathbf{P}}) \approx (\mathbf{U}_o + \mathbf{U}_s \Delta)(\mathbf{U}_o + \mathbf{U}_s \Delta)^H \quad (22)$$

where Δ is given by

$$\Delta = -\Sigma_s^{-1} \mathbf{V}_s^H \mathbf{W}^H \mathbf{U}_o \quad (23)$$

GENERATION OF AN SIRV

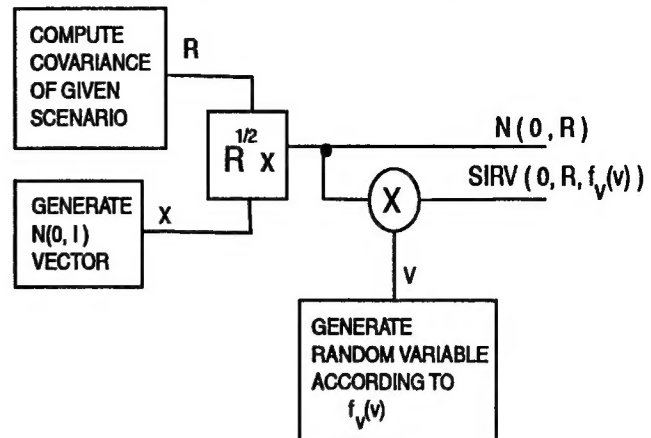


Figure 1: Block Diagram for Generation of an SIRV

4. SIMULATIONS

4.1. Noise Generation

The generation of the Gaussian and non-Gaussian noise is performed as shown in Figure 1. An independent realization of a Gaussian vector with independent elements is generated and then colored based upon the covariance of the chosen scenario. To produce the SIRV, an independent realization is drawn from the characteristic PDF and used to scale the random vector.

4.2. Simulation Steps

The steps of the simulation are illustrated in Figure 2. First, a set of K signal free independent vectors from the Gaussian and non-Gaussian interference scenarios are generated as described above. Note that the Gaussian and SIRV interference span the identical subspace. Background Gaussian noise is then added to these realizations after which a PCI weight vector is calculated for both the Gaussian and SIRV cases. For these simulations it was assumed that the rank of the interference was known. Another independent signal free Gaussian and non-Gaussian interference realization is generated with background noise added. The PCI weight vectors are applied to their respective interference types to compute an output sample under the H_0 assumption. A signal, αS is then added and weight vectors applied to obtain a sample under the H_1 condition. These steps are repeated for each trial of the simulation. False alarm thresholds are then set using percentages of the data generated under the H_0 assumption.

4.3. PCI Performance

For the purpose of simulation we consider a 16 element array with two jammers located at -10 and $+10$ degrees relative to broadside. The desired target is assumed to be at broadside. The jammers each have a power level of $+20dB$ relative to the background noise which is always white Gaus-

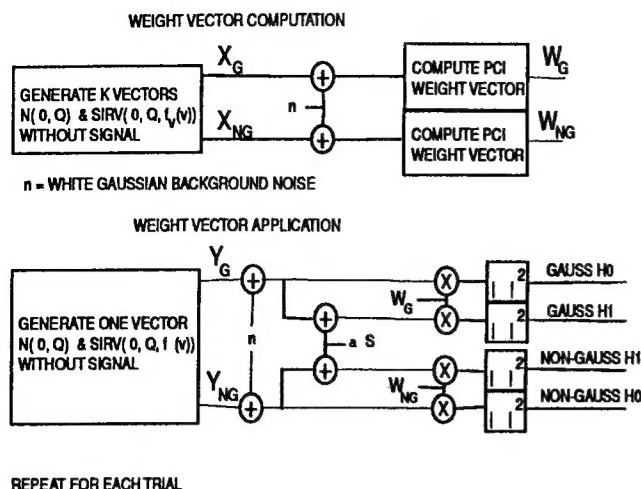


Figure 2: Block Diagram for PCI Simulation

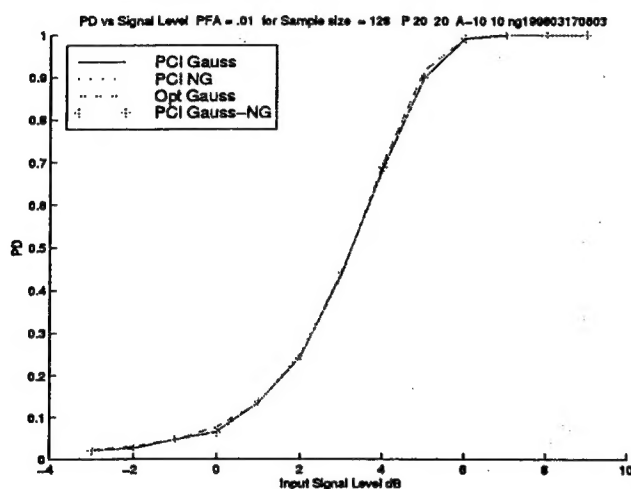


Figure 3: PD versus Signal Level for Gaussian and Non-Gaussian with Gamma Characteristic PDF, Sample Size = 128, $\alpha = .5$

sian. The jammer signals are taken to be independent, identically distributed Gaussian or Spherically Invariant Random vectors with the same covariance matrix. A Gamma characteristic PDF is used for the SIRV considered.

In the case of strong interference, the optimum Gaussian detector for known statistics is essentially a projection onto the null space of the interference. Since an SIRV with the same covariance matrix occupies the same subspace, this detector will have almost identical performance for the Gaussian and Non-Gaussian case. This is demonstrated in Figure 3 by letting the training sample size grow large and noting that all cases converge to one level of performance. Plotted is the probability of detection versus input signal level for PCI on Gaussian data (PCI Gauss) and PCI on non-Gaussian data (PCI NG) with a false alarm rate of .01 determined from the data generated under the H_0 assumption. Also, the optimum Gaussian performance (OPT Gauss) is plotted for comparison (PCI Gauss-NG will be discussed later).

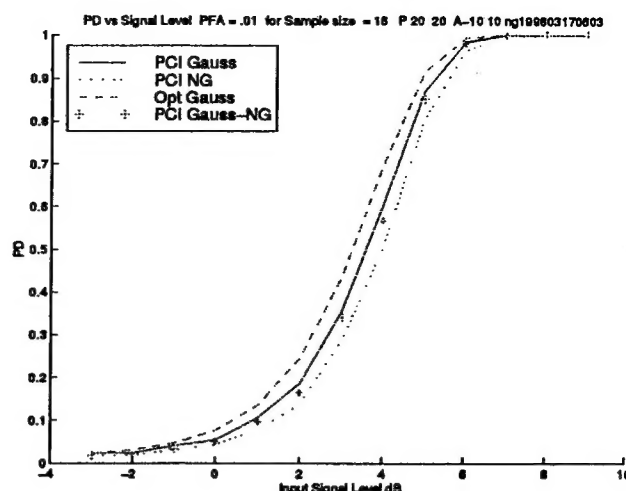


Figure 4: PD versus Signal Level for Gaussian and Non-Gaussian with Gamma Characteristic PDF, Sample Size = 16, $\alpha = .5$

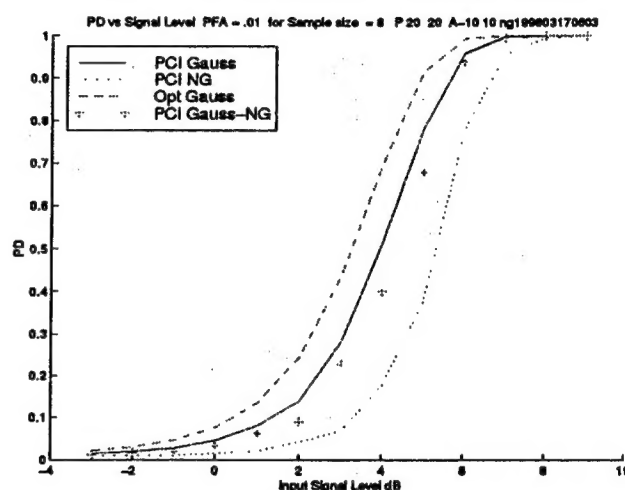


Figure 5: PD versus Signal Level for Gaussian and Non-Gaussian with Gamma Characteristic PDF, Sample Size = 8, $\alpha = .5$

Figure 4 shows the probability of detection versus input signal level for a smaller set of training samples. Sixteen signal free vector samples were used to compute an estimate of the PCI weight vector. With 16 samples for training, the performance of PCI is near the performance of the Gaussian detector of known statistics and there is only a small degradation for the non-Gaussian interference.

Figures 5 and 6 shows the PD curve for the case of 8 and 4 training samples. The performance of PCI on the Gaussian interference shows some degradation compared to the 16 sample training set. The performance of PCI on the non-Gaussian interference shows a more pronounced degradation than in the Gaussian case.

4.4. Discussion of Results

If the interference is strong and no assumptions are made about the strength of the signal then the performance of

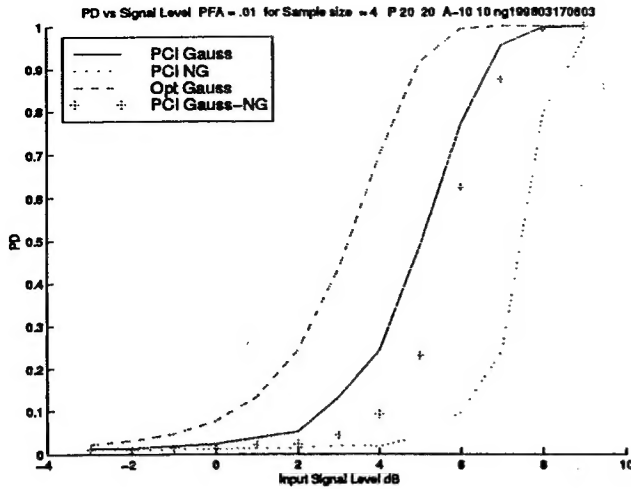


Figure 6: PD verses Signal Level for Gaussian and Non-Gaussian with Gamma Characteristic PDF, Sample Size = 4, $\alpha = .5$

using a projection will be identical for Gaussian or non-Gaussian interference. When the subspace is estimated, performance degrades as residual interference is allowed to pass through the projection matrix. The degradation will depend on the accuracy of the estimated subspace and the type of interference that is allowed to pass (Gaussian or non-Gaussian). Refer now to the PCI Gauss-NG points of Figure 5 and 6. These points were determined by computing the projection matrix using Gaussian interference training samples but then performing detection on the non-Gaussian interference. With moderate signal levels, the performance approaches that of the Gaussian case; the estimation of the projection matrix is the dominant factor in determining performance and the type of residual interference is less significant. For low signal levels, the more accurate projection estimation provides less of a performance increase.

The performance of PCI for moderate signal levels can be characterized by how well the subspace can be estimated in the presence of the different interference models. When the interference is Gaussian, the elements of Σ_s in equation 23 are distributed as a scaled χ^2_{2K} where K is the number of training samples. The terms W and U_o will be the same for Gaussian and non-Gaussian interference, and the V_s term will be statistically identical since each realization is independent. Therefore, the only difference lies in the distribution of the elements of Σ_s . Figures 7 and 8 show the histogram of the largest singular value for the non-Gaussian training data of the 128 and 4 sample cases respectively. Overlayed on each is the χ^2 distribution for Gaussian interference. The histograms are more spread than the Gaussian interference and for the smaller sample size the spread is skewed more to the smaller values. Also, note that the peak of the histogram is more left of the χ^2 overlay for the smaller sample size. Thus one can expect a higher percentage of larger perturbations with the non-Gaussian data than the Gaussian. It should be noted that there were no subspace swaps with either training size for the Gaussian interference

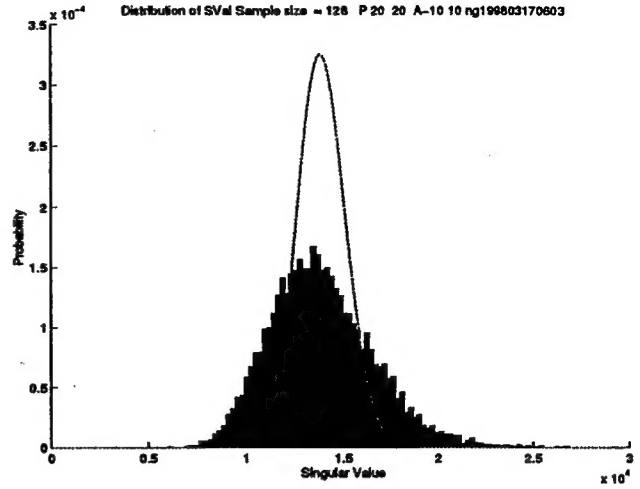


Figure 7: Distribution of Largest Singular Value Gamma Characteristic PDF, Sample Size = 16, $\alpha = .5$

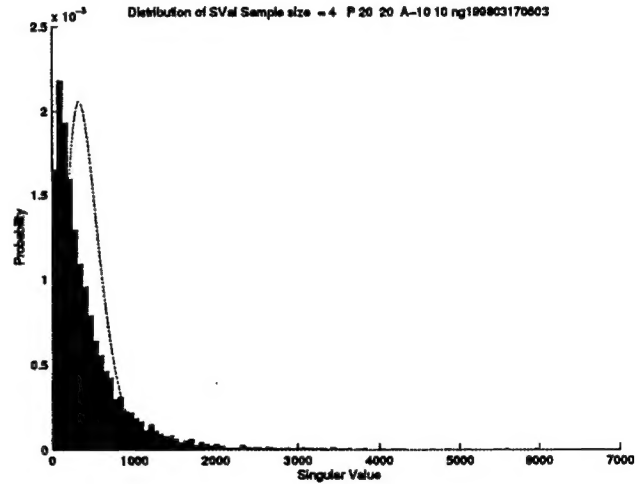


Figure 8: Distribution of Largest Singular Value Gamma Characteristic PDF, Sample Size = 8, $\alpha = .5$

and only a very small ($\sim 1\%$) number of swaps for the sample size of 4 with the non-Gaussian interference. This was not significant enough to effect the detection curves.

5. CONCLUSIONS

In the case of moderate to large sample size for secondary data, the Gaussian performance of the PCI method for strong low rank interference will be maintained for non-Gaussian SIRD interference. However, as the sample size decreases, the performance in non-Gaussian interference degrades faster than for Gaussian interference. For moderate to large signal levels, this degradation can be attributed to poorer estimation of the projection matrix which is the result of a larger spread in the distribution of the singular values of the secondary data matrix. In particular, the non-Gaussian singular values are more likely to be smaller than those from Gaussian data. The extent to which this is true will be a function of the characteristic PDF used to describe the non-

Gaussian SIRV. For lower signal levels, the non-Gaussian nature of the residual interference limits the performance. Future work should include possible remedies for robust estimation of the projection matrix when the interference is may be non-Gaussian.

6. REFERENCES

- [1] L.E. Brennan and I.S. Reed. Theory of adaptive radar. *IEEE Transactions on Aerospace and Electronic Systems*, AES-9(2):237-252, March 1973.
- [2] L.W. Brooks and I.S. Reed. Equivalence of the likelihood ratio processor, the maximum signal-to-noise ratio filter, and the wiener filter. *IEEE Transactions on Aerospace and Electronic Systems*, pages 690-692, Septmeber 1972.
- [3] Alfons J. Claus, T.T. Kadota, and Dennis M. Romain. Efficient approximation of a family of noises for application in adaptive spatial processing for signal detection. *IEEE Transactions on Information Theory*, IT-26(5):588-595, September 1980.
- [4] I. KIRSTEINS and D. TUFTS. Adaptive detection using low rank approximation to a data matrix. *IEEE Transactions on Aerospace and Electronic Systems*, 30(1):55-67, January 1994.
- [5] I.P. KIRSTEINS and D.W. TUFTS. On the probability density of signal-to-noise ratio in an improved adaptive detector. In *IEEE Proc ICASSP-85*, pages 572-575, Tampa, FL, April 1985.
- [6] I.P. KIRSTEINS and D.W. TUFTS. Rapidly adaptive nulling of interference. In M. Bouvet and G. Bienvenu, editors, *High Resolution Methods in Underwater Acoustics*. Springer-Verlag, New York, N.Y., 1991.
- [7] B. PICINBONO. On circularity. *IEEE Transactions on Signal Processing*, 42:3473-3482, 1994.
- [8] Muralidhar Rangaswamy, James H. Michels, and Donald D. Weiner. Multichannel detection for correlated non-gaussian random processes based on innovations. *IEEE Transactions on Signal Processing*, 43(8):1915-1922, 1995.
- [9] Muralidhar Rangaswamy, Donald D. Weiner, and Aydin Ozturk. Non-gaussian random vector identification using spherically invariant random processes. *IEEE Transactions on Aerospace and Electronic Systems*, 29(1):111-124, 1993.
- [10] Muralidhar Rangaswamy, Donald D. Weiner, and Aydin Ozturk. Computer generation of correlated non-gaussian radar clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 31(1):106-116, 1995.
- [11] J.K. Thomas, L.L. Scharf, and D.W. Tufts. The probability of a subspace swap in the svd. *IEEE Transactions on Signal Processing*, 43(3):730, March 1995.
- [12] D.W. Tufts, I. KIRSTEINS, and R. KUMARESAN. Data adaptive detection of a weak signal. *IEEE Transactions on Aerospace and Electronic Systems*, AES-19(2):313-316, March 1983.
- [13] D.W. Tufts, A.C. Kot, and R.J. Vaccaro. The analysis of threshold behavior of svd-based algorithms. In *Proceedings of the 21st Asilomar Conference on Signals, Systems and Computers*, pages 550-554, Pacific Grove, CA, November 1987.
- [14] R.J. Vaccaro. A second order perturbation expansion for the svd. *SIAM J. Matrix Anal. Appl.*, 15(2):661-671, April 1994.
- [15] K. Yao. A representation theorem and its application to spherically invariant random processes. *IEEE Transactions on Information Theory*, IT-19:600-608, 1973.